

STIC-Biotech/ChemLib

97802

From: Whiteman, Brian
Sent: Friday, April 18, 2003 10:19 AM
To: STIC-Biotech/ChemLib
Subject: seq search

STIC-Biotech/ChemLib
APR 18 2003
10:19 AM

09/729,264 11/26/00 Welcher et al.

search seq id nos 1, 3, and 5 against us patent and us patent application databases.

search seq id nos 2, 4, and 6 against us patent and us patent application databases.

Thanks,
Brian Whiteman, 11e12
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
Crystal Mall 1, 11A16
(703) 305-0775

Mary Jane Ruhl
Tech. Info. Specialist, STIC
TC-1600
CM-1, Room 6A-06
Phone: 605-1155

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 4/30/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Wed Apr 30 10:00:21 2003

us-09-729-264-2.rsp1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:24:32 ; Search time 36.7435 seconds
(without alignments)
2142.147 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059 1 MGIVFLHSGSGNEVIEGP.....HPQAFNLASPEKYSNTTV 382

Sequence: BLOSUM62

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_dacteryap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1694	82.3	315	4 Q9NS15	Q9NS15 homo sapien
2	575.5	28.0	270	11 Q9DP82	Q9DP82 mus musculu
3	141	168	11	Q9DP92	Q9DP92 mus musculu
4	136.5	6.6	173	11 Q9D4K2	Q9D4K2 mus sapien
5	135	6.6	1496	4 Q92626	Q92626 xenopus lae
6	133	6.5	1427	13 Q91562	Q91562 gallus gall
7	132.5	6.4	344	13 Q9DF61	Q9DF61 mus musculu
8	131	6.4	697	11 Q9DBP0	Q9DBP0 mus musculu
9	131	6.4	697	11 Q92290	Q92290 mus musculu
10	129.5	6.3	315	13 Q9DP15	Q9DP15 gallus gall
11	127.5	6.2	315	13 Q9DP15	Q9DP15 mus musculu
12	127.5	6.2	344	13 Q93242	Q93242 gallus gall
13	127.5	6.2	871	11 Q61987	Q61987 mus musculu
14	127.5	6.2	881	11 Q61988	Q61988 mus musculu
15	127.5	6.2	197	5 Q17641	Q17641 caenorhabd
16	127				

17	127	6.2	1445	11	Q63155	Q63155 homo sapien
18	126.5	6.1	869	4	Q15146	Q15146 gallus gall
19	126	6.1	313	13	Q57596	Q57596 caenorhabd
20	125.5	6.1	166	5	Q950Y1	Q950Y1 caenorhabd
21	125.5	6.1	6632	5	Q01761	Q01761 homo sapien
22	123.5	6.0	344	4	Q9P121	Q9P121 homo sapien
23	123.5	6.0	6620	4	Q96AA2	Q96AA2 caenorhabd
24	122.5	5.9	188	5	Q18238	Q18238 homo sapien
25	122.5	5.9	455	4	Q9D1R0	Q9D1R0 drosophila
26	120	5.8	959	5	Q9N9Y9	Q9N9Y9 drosophila
27	120	5.8	968	5	Q9W4T9	Q9W4T9 drosophila
28	119	5.8	975	5	Q9J174	Q9J174 drosophila
29	118	5.7	695	11	Q9J309	Q9J309 rattus norv
30	117.5	5.7	275	11	Q9H1X9	Q9H1X9 mus musculus
31	117.5	5.7	390	4	Q9H1X9	Q9H1X9 homo sapien
32	117.5	5.7	5636	4	Q96RW7	Q96RW7 homo sapien
33	117	5.7	325	6	Q02838	Q02838 sus scrofa
34	116.5	5.7	345	5	Q9W4U1	Q9W4U1 drosophila
35	116	5.6	2221	5	Q9H1M1	Q9H1M1 drosophila
36	116	5.6	4370	4	Q9H3V5	Q9H3V5 homo sapien
37	115	5.6	6632	5	Q17362	Q17362 caenorhabd
38	115.5	5.6	196	4	Q9H0T1	Q9H0T1 homo sapien
39	114.5	5.6	152	5	Q9YVX3	Q9YVX3 caenorhabd
40	114	5.5	46	4	Q16861	Q16861 homo sapien
41	114	5.5	327	4	Q96107	Q96107 homo sapien
42	114	5.5	569	11	Q920G3	Q920G3 mus musculus
43	114	5.5	1395	5	Q44924	Q44924 drosophila
44	114	5.5	1395	5	Q9W213	Q9W213 drosophila
45	114	5.5	2673	4	Q96SC3	Q96SC3 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	315 AA.
Q9NS15	Q9NS15	Q9NS15		
AC	01-OCT-2000 (TREMURel. 15, Created)			
DT	01-OCT-2000 (TREMURel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMURel. 19, Last annotation update)			
DE	IGSF5 protein (Fragment).			
GN	IGSF5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,			
RA	Toyoda A., Ishii K., Tokoi Y., Choi D.K., Soeda E., Ohki M.,			
RA	Takegi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,			
RA	Wenzel U., Delabar J., Rumpf K., Lehmann R., Patterson D.,			
RA	Kelchwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,			
RA	Rosenblatt A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,			
RA	Shintani A., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,			
RA	Minoshima S., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,			
RA	Scharfe M., Schenck A., Klages S., Hennig S., Riesselmann L., Dagand E.,			
RA	Kammer J., Beck A., Klages S., Gardiner K., Nizetic D., Francis F.,			
RA	Wehmeyer S., Borzym K., Yaspo M.L.,			
RA	Lehrach H., Reinhardt R., Yaspo M.L.,			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL163280; CAB90447.1;			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR003599; Ig.			
DR	InterPro: IPR003066; Ig-MHC.			
DR	Pfam: PF00047; Ig_2.			
DR	SMART: SM00409; Ig_1; UNKNOWN_1.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
FT	NON_TER			
FT	NON_TER			
FT	SEQUENCE			
SO	SEQUENCE			

Query Match

Best Local Similarity 99.7%; Score 1694; DB 4; Length 315;
Matches 314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 SGGSENYIEGPQNAVYLGKSGARFNCYVSGKRLMALSDMVLVSRPMEPIITNRT 69
DB 1 SGGSENYIEGPQNAVYLGKSGARFNCYVSGKRLMALSDMVLVSRPMEPIITNRT 69
QY 70 SGRYDGGNFTSEMIHNEPDSGNIRCSLONSLRHSAYLVYVWGEFISVNLVYA 129
DB 61 SGRYDGGNFTSEMIHNEPDSGNIRCSLONSLRHSAYLVYVWGEFISVNLVYA 129
QY 130 ENPEPEVCLPSHWRTRLPDISWELGLVSHSSYFVEPEPSDLOSANVILALPQSNGLT 180
DB 121 ENPEPEVCLPSHWRTRLPDISWELGLVSHSSYFVEPEPSDLOSANVILALPQSNGLT 180
QY 190 CVATMKSILKARKSATVNLVYIRCPDPTGGGINIPGVLSLPSLGFSLPTWKGVLGLAGT 240
DB 181 CVATMKSILKARKSATVNLVYIRCPDPTGGGINIPGVLSLPSLGFSLPTWKGVLGLAGT 240
QY 250 MLTPPTCTLTIRCCCRRCRCCGCCGCCFCCRRKRGFRIOFQKSEKETEETES 309
DB 241 MLTPPTCTLTIRCCCRRCRCCGCCGCCFCCRRKRGFRIOFQKSEKETEETES 309
QY 310 GNEGSGYNSDEQKT 324
DB 301 GNEGSGYNSDEQKT 315

RESULT 2

09D8G2 ID 09D8G2 PRELIMINARY; PRT; 270 AA.
AC 09D8G2
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 2010003D20RIK protein (RIKEN CDNA 2010003D20 gene).
GN 2010003D20RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
RA MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Casavant T.,
RA Schirml L.M., Staubli F., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Blake J., Offield T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Hill D., Hofmann M., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Watanabe-Borja A., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RP
RA
RA SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008060; BAB25436.1;
DR MGD; BC004806; AAH04806.1;
MCD; MGI:1919308; 2010003D20RIK.

DR InterPro: IPR003599, Ig.
DR InterPro: IPR003006, Ig_MHC.
DR SMART: SM00409; Ig_1.
DR PROSITE: PS00290; Ig_MHC; UNKNOWN.1.
SQ SEQUENCE 270 AA; 29604 MW; A339C273DA950DPE CRC64.

Query Match 28.0%; Score 575.5; DB 11; Length 270;
Best Local Similarity 35.1%; Pred. No. 2.3e-44;
Matches 136; Conservative 42; Mismatches 73; Indels 137; Gaps 6;

QY 3 LVIF--LHSGGSENYIEGPQNAVYLGKSGARFNCYVSGKRLMALSDMVLVSRPME 60
DB 12 LVILAQLTASGSSYQIIEGPQNAVYLVLDSEHFNCTYHGMKLMWLNQVYLSITTG 71
QY 61 PIITNRTSRYDGGNFTSEMIHNEPDSGNIRCSLONSLRHSAYLVYVWGEFISVNL 120
DB 72 PIITNRTSRYDGGNFTSEMIHNEPDSGNIRCSLONSLRHSAYLVYVWGEFISVNL 120
QY 121 IPSVNLVVAENPEVCLPSHWRTRLPDISWELGLVSHSSYFVEPEPSDLOSANVILAL 180
DB 126
QY 181 TPQSNGLTFCVATMKSILKARKSATVNLVYIRCPDPTGGGINIPGVLSLPSLGFSLPTWG 240
DB 126
QY 241 KVLGLAGTMLTPPTCTLTIRCCCRRCRCCGCCGCCFCCRRKRGFRIOFQKSEKETEETES 297
DB 140 ILLAVAFSLILILIVLIIIFCC-----DSIGEEGPALPTVA 139
QY 298 KEKTKN---EFTEGSGNYSYNSDEQKTDTASLPKSCSSDPORSSCCPPHQRAD 184
DB 185 NMRTKADPEPKIKSGKNGYSSDEKAAQTASLPKSAFVSLPKKSSSL--PYQELN 242
QY 355 QRPPEPASHPQASNLASPEKYSNTTVV 382
DB 243 KHQGPATHRFVSPDIASPKVKNVTLV 270

RESULT 3

09D912 ID 09D912 PRELIMINARY; PRT; 168 AA.
AC 09D912
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 4931420D14RIK protein.
GN 4931420D14RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RA MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Casavant T.,
RA Schirml L.M., Staubli F., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Blake J., Offield T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Hill D., Hofmann M., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Watanabe-Borja A., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RP
RA
RA SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008060; BAB25436.1;
DR MGD; BC004806; AAH04806.1;
MCD; MGI:1919308; 2010003D20RIK.

FT NON_TER 1 1
 SQ SEQUENCE 1496 AA; 167209 MW; E9B9A7069BFLABEF CRC64;

Query Match
 Best Local Similarity 25.0%; Score 135; DB 4; Length 1496;
 Matches 73; Conservative 36; Mismatches 127; Indels 56; Gaps 15;

QY 9 GSGGNEVIEGPNATVYKGSQAFNCTVS--QCGKLIMALSDMVLSVRPEPIITNDR 67
 Db 354 GSPARPEFVIOPTQEVILVAGESVTLBESATGHPPIISWTGRIPPLPDPFVNITPS-- 411
 QY 68 FTSORYOGGNTFSEMIHNVPESDSGNIRCSLONS--RLHGSAYLTVQMGELFLPSVN 125
 Db 412 -----GG-----LTIQNVVQDSDGEYACASATNNIDSVHATFLIYQALQFIVTQD 458
 QY 126 LVVAENPECEVTCLPSHMTRLPDISW--ELGLVSHSSYFVEPEPDSQANSIATLPQS 184
 Db 459 RVVIEGQTVDFQC--EAKGNPPPIYIAWTKGSLSDVDRHLVLSGTLR--ISGVALHDO- 514
 QY 185 NGTLTCVATWKSARKSATVNLTV-----INCPQDT---GGGINIPGVLSLPSL 232
 Db 515 -GGYECQAV--NIGSQKVAHLTVQPRVTVFASIPSDTVEVGANYQLP-----CSSQ 566
 QY 233 GFSLP--TWGKVGGLT--AGTMLTPCTCTLT-----RCCCGRRRCG 271
 Db 567 GEPERATYMNKDGVOVTESGKFHISPEGFLITINDGPADAGRYCVARNITG 618

RESULT 6

Q91562 PRELIMINARY; PRT; 1427 AA.

AC Q91562;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN XCCA.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95113183; PubMed=7813784;
 RA Pierceall W.E.; Reale M.A.; Candia A.F.; Wright C.V.; Cho K.R.;
 RA Reardon E.R.;
 RT "Expression of a homologue of the deleted in colorectal cancer (DCC)
 gene in the nervous system of developing Xenopus embryos."
 RL Dev. Biol. 166:654-665(1994).
 DR EMBL: U10986; AAA70168.1; -.
 DR HSSP: P40189; 180V.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00041; fn3; 6.
 DR PRINTS: PR00014; ENTPEPIT.
 DR SMART: SM00060; FN3; 6.
 DR SMART: SM00408; IgC2; 3.
 DR SMART: SM00410; Ig_Like; 2.
 KW Immunoglobulin domain; Repeat.
 SQ SEQUENCE 1427 AA; 156533 MW; 61FEA12C8A674972 CRC64;

Query Match
 Best Local Similarity 25.7%; Score 133; DB 13; Length 1427;
 Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 17 IESPNATVYKGSQAFNCTVSQGW--KLIMALSDMVLSVRPEPIITNDRFSORYD 74
 Db 243 LQPSVVAIEGQDAVLECAVS--GYPTPIYMWQD-----EPVPIRTR---KYS 288

QY 75 QGNETSEMIHNVPESDSGNIRC--SLONSRLHGSAYLTVQMGELFLPSVNLVAENE 132
 Db 289 VLGG--SNLISNTVDDACATYCAVAKNENTSFSADLTVMPPQFLNHPANLVAYESM 346
 QY 133 PCEVTC-----PSHWRRLPDISW--ELGLVSHSSYFVEPEPDSQANSIATLPQSNGT 187
 Db 347 DIEECVASKPS-----PTVKWTKNGEVLPISDFQIYDGSNLR-----ILGLVSDGXY 397
 QY 188 LTCVA-----TWKSLARKSATVNLTV--RCQDPTGGGINIPGVLS 228
 Db 398 YOCIAENENIGTVAQLIIPDPAYPSSILPSADRV-----VPLVSS 442

RESULT 7

Q9DF61 PRELIMINARY; PRT; 344 AA.

AC Q9DF61;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN OBGM.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-BRAIN.
 RX MEDLINE=20499204; PubMed=11042360;
 RA Lodge A.P.; Howard M.R.; McNamee C.J.; Moss D.J.;
 RT "Co-localisation, heterophilic interactions and regulated expression
 of Irgon family proteins in the chick nervous system."
 RL Brain Res. Mol. Brain Res. 82:84-94(2000).
 DR EMBL: AF292934; AAG01877.1; -.
 DR InterPro: IPR003599; Ig_C.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00409; Ig; 3.
 DR SMART: SM00408; IgC2; 3.
 DR SMART: SM00410; Ig_Like; 2.
 KW Immunoglobulin domain.
 SQ SEQUENCE 344 AA; 37531 MW; 37FE6051CBFE0E7B4 CRC64;

Query Match
 Best Local Similarity 25.5%; Score 132.5; DB 13; Length 344;
 Matches 70; Conservative 38; Mismatches 104; Indels 63; Gaps 16;

QY 4 VIFLHSG---SGNEVI--EGPNATVYKGSQAFNCTVSQCGKLIMALSDMVVL---- 54
 Db 21 LFLVPAAGVPRSGDATFPKAMDVYTRQESATLCTYDDRVRRVAV--LNSSTILYAGN 79
 QY 55 ---SVRPEPIITNDRFSORYOGGNTFSEMIHNVPESDSGNIRCSQ-----NSRL 105
 Db 80 DKMSIDNRVILSN---TKTQY-----SIKIHVDYDGGPTCSQVDNHNKRTSRV 128
 QY 106 HGSAYLTVQMGELFLPSVNLVAENPECEVCL-----PSHWRRLPDISWELGLVSHSS 161
 Db 129 H---LIVQVPQIVNISSDIYNESSSTYLWCLAFGRE-----PVTWR---HLSGKG 176
 QY 162 YFVPEPDSQANSIATLPQSNGTLTCVATWKSARKSATVNLTV-----TRCPQDT 216
 Db 177 QGSVSEDEYLE---ITGTREOSGEYECNAV--NDVAVPDRKKVYVYNNPPYISNAKNT 231
 QY 217 GGGINIPGVL-----SISLSLGSFLPTWKGKVGUGIA 247
 Db 232 GASVCGKIGITLQCEASAVVAERQ---WFKEDTRLA 263

[illegible]

QY 4 VIFLHSG-----SGNEVI-EGPONAIVLKSGSARFNCTVSGMKLIMVLSDMVIVSVRP 58
 DB 21 LFLVPTGVPVRSRSDATPPKAMDNVTVRQGESATLRCTIDNRVRYVAMLNRSIT-----74
 QY 59 MEPIITNDNF-TSORYDQGNFTSE--MIINVEPSDGNIRCSIQ-----NSRLHGA 109
 DB 75 --PYAGNDKWCIDPRVLLIGNTQTOYSIEIONVDVYDEGPTYSQVQIDNHFKTSRVH---129
 QY 110 YLVQVYMGELFIPSVNLVVAENEPCEVCLPSSHMTRLPDISWELGLVSHSSYYFVPEPS 169
 DB 130 -LIVQVSPKTVELISDSISINEGNISLTICIAGRE-PLVYWR---HISPKAVGFSDE 184
 QY 170 DLQSAVSIILATPQSGTTLVCAVTKMSLKARSAVNLTVICP-----QDTGGGINIPG 224
 DB 185 YLQ-----IGITREOSGEYECSSAS-NDVAAPVPRVRYVNPYPYSEAKGTGVVQKQ 239
 QY 225 VL-----SLPSLGFSLPTWCK 241
 DB 240 TLQCEASAVPSAEFO---WCK 257

RESULT 13

093242 PRELIMINARY; PRT; 344 AA.
 AC 093242; 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CEPU-1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kikura Y., Shirabe K., Fukushima M., Takeshita M., Tanaka H.;
 RT "CEPU-1: an immunoglobulin superfamily molecule, has cell adhesion
 RT activity and shows dynamic expression patterns in chick embryonic
 RT spinal cord";
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB011810; BRA31514.1.
 DR InterPro: IPR003398; I9_C2.
 DR InterPro: IPR003600; I9_Like.
 DR InterPro: IPR003006; I9_MHC.
 DR Pfam: PF00047; I9_3.
 DR SMART: SM00408; Igc2; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR Immunoglobulin domain.
 KW SEQUENCE 344 AA; 37613 MW; 22CAAF526A857E CRC64;

Query Match 6.28; Score 127.5; DB 13; Length 344;
 Best Local Similarity 25.44; Pred. No. 0.0023; 98; Indels 61; Gaps 15;
 Matches 68; Conservative 41; Mismatches 98;

QY 4 VIFLHSG-----SGNEVI-EGPONAIVLKSGSARFNCTVSGMKLIMVLSDMVIVSVRP 58
 DB 21 LFLVPTGVPVRSRSDATPPKAMDNVTVRQGESATLRCTIDNRVRYVAMLNRSIT-----74
 QY 59 MEPIITNDNF-TSORYDQGNFTSE--MIINVEPSDGNIRCSIQ-----NSRLHGA 109
 DB 75 --PYAGNDKWCIDPRVLLIGNTQTOYSIEIONVDVYDEGPTYSQVQIDNHFKTSRVH---129
 QY 110 YLVQVYMGELFIPSVNLVVAENEPCEVCLPSSHMTRLPDISWELGLVSHSSYYFVPEPS 169
 DB 130 -LIVQVSPKTVELISDSISINEGNISLTICIAGRE-PLVYWR---HISPKAVGFSDE 184
 QY 170 DLQSAVSIILATPQSGTTLVCAVTKMSLKARSAVNLTVICP-----QDTGGGINIPG 224
 DB 185 YLQ-----IGITREOSGEYECSSAS-NDVAAPVPRVRYVNPYPYSEAKGTGVVQKQ 239
 QY 225 VL-----SLPSLGFSLPTWCK 241
 DB 240 TLQCEASAVPSAEFO---WCK 257

DB 233 VVVGQGIIMCEASAVPSADFQ---WCK 257

RESULT 14

061987 PRELIMINARY; PRT; 871 AA.
 AC 061987; 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Nsk2 protein precursor.
 GN MUSK OR NSK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MYOBLAST;
 RX MEDLINE=9534951; PubMed=7624144;
 RA Ganju P., Nails E., Brennan J., Reith A.D.;
 RT "Cloning and developmental expression of Nsk2, a novel receptor
 RT tyrosine kinase implicated in skeletal myogenesis.";
 RL Oncogene 11:281-290(1995).
 CC -1-SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; X86444; CAA60165.1.
 DR HSP; P1362; IFK.
 DR MCD; MGI:103581; Mus.
 DR InterPro: IPR002453; Beta.tubulin.
 DR InterPro: IPR000719; Euk.pkinase.
 DR InterPro: IPR000024; Fz_domain.
 DR InterPro: IPR003598; I9_C2.
 DR InterPro: IPR003600; I9_Like.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00392; Fz; 1.
 DR Pfam; PF00047; I9_3.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000001; Euk_pkinase; 1.
 DR SMART; SM00408; Igc2; 2.
 DR SMART; SM00410; Ig_Like; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS50038; Fz; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
 DR APB-binding; Immunoglobulin domain; Serine/threonine-protein kinase;
 KW Signal; Transferrase.
 FT SIGNAL 1
 FT SEQUENCE 871 AA; 97047 MW; F3C53D6A6E702AB CRC64;

Query Match 6.28; Score 127.5; DB 11; Length 871;
 Best Local Similarity 21.78; Pred. No. 0.0077;
 Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

QY 15 EVIEGPNATVVLKSGSARFNCTVSGMKLIMVLSDMVIVSVRPMEPIITNDNRSORY 73
 DB 122 KITRPINVKIIEGLKAVLPCTMGNRPSPYSWTKG-----NALNENRIAALE 171
 QY 74 DOGQNFTEMIHNEVPSDGNIRCSLNSRLHGSAY-----LTVQVYMGELFIPSVNLV 128
 DB 172 -----SGSLRHNVQKEDAGYRCVAKNSL--GTAISKLVKLEVEVIGRILAPESHNV 223
 QY 129 AENEPEVTC-----LPSSHMTRLPDISW-ELGLVSHSSYYVPEPSDLOSASVSLATPQ 183
 DB 224 TEGSEFVLTCTELGIP-----VPTISWENGNAVSSGSIQESVORVIDSRLQFLTKP- 277
 QY 184 SNGTLCVAT-----WKSLSKARSAVNLTVIR-----CPDPTGGGINIFG-----V 225

Db 278 --GLYTCIATNKHGEKSTAKAATVSIAMSKSQKDSQGYCAQYEGEVLMOGPEKML 335
 QY 226 LSLPSLGFSLP-----TWGKVGGLAGTMTLPTCTLTTRCCRRCCGCN-- 273
 Db 336 LVFLPTSHRDEDAOELLHTAMNEL-----KAVSPICRPAEALLCYHLFLCCSPG 388
 QY 274 -----CCRCRC-----FCCR 283
 Db 389 VVPTPMPICREYCIAMKELFCAR 411

RESULT 15

Q61988

ID 061988 PRELIMINARY; PR; 881 AA.

AC 061988;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

GN MUSK OR NSK2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OY NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=MYOBLAST;

RX MEDLINE=9534951; PubMed=762144;

RA Ganju P., Walls E., Brennan J., Reith A.D.;

RT "Cloning and developmental expression of Nsk2, a novel receptor

RT tyrosine kinase implicated in skeletal myogenesis,"

RL Oncogene 11:281-290(1995).

CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: X86445; CAA60166.1; -

DR HSP; P11362; IFGK.

DR MGI:103581; Mus.

DR InterPro: IPR002453; Beta-tubulin.

DR InterPro: IPR000719; Euk_pkinase.

DR InterPro: IPR000024; Fz_domain.

DR InterPro: IPR003598; Ig_c2.

DR InterPro: IPR003600; Ig_1like.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR002290; Ser_thr_pkinase.

DR pfam: PF01392; Fz; 1.

DR pfam: PF00047; Ig; 3.

DR pfam: PF00069; pkinase; 1.

DR PRINTS: PR00109; TYRKINASE.

DR PRODOM: PD000001; Euk_pkinase; 1.

DR SMART: SM00408; IG_c2; 2.

DR SMART: SM00410; IG_1like; 1.

DR SMART: SM00219; Tyrc; 1.

DR PROSITE: PS50038; Fz; 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.

DR ATP-binding; Immunoglobulin domain; Serine/threonine-protein kinase;

KW Signal; Transferase; 21

FT SIGNAL 1

SQ SEQUENCE 881 AA; 96435 MW; EA0D0282E9B28ED7 CRC64;

Query Match 6.2%; Score 127.5; DB 11; Length 881;

Best Local Similarity 21.7%; Pred. No. 0.0078;

Matches 70; Conservative 41; Mismatches 135; Indels 87; Gaps 16;

QY 15 EVIEGPNATVLCGSAARECTYSQGW-LIMWALSDMVLVSVRPMPEPTITNDRTSORY 73

Db 122 KIRPPTLVNKTIGLAAVLPCTMGNPKPSVMIKGD-----NAKRENSRIALE- 171

QY 74 DQGNFTSEMIHNVEPSDSGNIRCSLRHGSAY-----LTVQVNGELFIPSVNLV 128

Db 172 -----SGSLRHHVQKEDAGQYCAVAKNSL--GTAYSKLVKLEVEVIGRIIRAPESHNV 223
 QY 129 AENEPEVTC-----LPSHWTRLPDISW-ELGILVSHSYTFVPEPSDIQSAVSIATLPQ 183
 Db 224 TFGSEVTLRCTELGIP-----VPTISWINGNAVSSGSIQESVKRVIDSRQLQFLTKP- 277
 QY 184 SNGTLICVAT-----WKSIAKRSATVNLVIR-----CPDGTGGGNIIPG-----V 225
 Db 278 --GLYTCIATNKHGEKSTAKAATVSIAMSKSQKDSQGYCAQYEGEVLMOGPEKML 335
 QY 226 LSLPSLGFSLP-----TWGKVGGLAGTMTLPTCTLTTRCCRRCCGCN-- 273
 Db 336 LVFLPTSHRDEDAOELLHTAMNEL-----KAVSPICRPAEALLCYHLFLCCSPG 388
 QY 274 -----CCRCRC-----FCCR 283
 Db 389 VVPTPMPICREYCIAMKELFCAR 411

Search completed: April 28, 2003, 21:07:34
 Job time : 42.7435 secs

Amgen, Inc. (US)
 Location/Qualifiers
 1.1168
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 8.1168
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28217.1"
 /db_xref="GI:19575329"
 /translation="MAYAGMENRDPGSSGSENYIEGPONAVLKGSOARFNCYVSG
 KMLIMALSDVAVLSVREMEPIITNDRETSQYDGGNTSEMIITHNVEPSDGINRC
 LONSRLHGSALIVQVMELEIPSNVLVAENNECEVCLPSHMTWLPDISMEGL
 VSHSYEVPEPSDLOSASVILATPOSNGTLCAVATKSKARSATVNLVIRCPD
 DTGGINIPGVLSLPSLGEFLPTWKGVLGAGMLLPTCTLTIRCCCRRCGCC
 NCCRCRCRRKRGFRIOFOKSKSEKTKNETETESGNGSYNSDEKTDITSLP
 PKSCSSPDEQNSCCGPHQADRPAPRPSHPQASFNILASPERVNTIV"

BASE COUNT 299 a 302 c 288 g 279 t
 ORIGIN

Query Match 100.0%; Score 1168; DB 6; Length 1168;
 Best Local Similarity 100.0%; Pred. No. 2.4e-270;
 Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGATCATGTGGCGAGGCCATGGAATAAGACCCACCGGTTGTGGCTGTAA 60
 DB 1 AGTGATCATGTGGCGAGGCCATGGAATAAGACCCACCGGTTGTGGCTGTAA 60
 QY 61 TGAAGTCATAGAAGGCCCAAAATGCAAGAGCTCGAAGGGCTCCAGGCTGCTCA 120
 DB 61 TGAAGTCATAGAAGGCCCAAAATGCAAGAGCTCGAAGGGCTCCAGGCTGCTCA 120
 QY 121 CTGCACCGTCTCCAGAGGCTGGAAGCTCATGTGGGCTCTCACTGACATGGTGTCT 180
 DB 121 CTGCACCGTCTCCAGAGGCTGGAAGCTCATGTGGGCTCTCACTGACATGGTGTCT 180
 QY 181 AAGCGTCAGAGCCCATGAGACCCCATCATCAACCAATGACCCCTTCACTCAAGAGTACCA 240
 DB 181 AAGCGTCAGAGCCCATGAGACCCCATCATCAACCAATGACCCCTTCACTCAAGAGTACCA 240
 QY 241 CCAGGCGGGGAATCTACCTCGAGATGATCATCCACAATGTGAGCCAGTGGATCGGG 300
 DB 241 CCAGGCGGGGAATCTACCTCGAGATGATCATCCACAATGTGAGCCAGTGGATCGGG 300
 QY 301 GAACATCAGATCAGCCTTCAGAAAGAGTGGCTGATGATGCTTACCTTACCGTCA 360
 DB 301 GAACATCAGATCAGCCTTCAGAAAGAGTGGCTGATGATGCTTACCTTACCGTCA 360
 QY 361 AGTTATGGAGAGCTGTTCATCCAGTGTATCTTGTAGTGGCTGGAATGAACCTTG 420
 DB 361 AGTTATGGAGAGCTGTTCATCCAGTGTATCTTGTAGTGGCTGGAATGAACCTTG 420
 QY 421 TGAAGTCTTGTCTACCTTCACACTGAGACCTGGCTCCGGATATTTCTCGGAGCTCG 480
 DB 421 TGAAGTCTTGTCTACCTTCACACTGAGACCTGGCTCCGGATATTTCTCGGAGCTCG 480
 QY 481 TCTCTGTGAGCCATTAAGTATTTTGTCCGAGCCAGGACCTTCAAAAGTGC 540
 DB 481 TCTCTGTGAGCCATTAAGTATTTTGTCCGAGCCAGGACCTTCAAAAGTGC 540
 QY 541 AGTGAGCATCTGGCTGAGCCACAGAGCAATGGACTTGAATTCGATGGCTGCTG 600
 DB 541 AGTGAGCATCTGGCTGAGCCACAGAGCAATGGACTTGAATTCGATGGCTGCTG 600
 QY 601 GAAAGAGCTGAAGCCCGCAAGTCTGCAAGTGAATCTCACTGTGATGGTGTCCCA 660
 DB 601 GAAAGAGCTGAAGCCCGCAAGTCTGCAAGTGAATCTCACTGTGATGGTGTCCCA 660
 QY 661 AGACACTGAGAGTGGTATTAATTCAGAGTGTATTAATTCAGAGTGTAAAGTTT 720
 DB 661 AGACACTGAGAGTGGTATTAATTCAGAGTGTATTAATTCAGAGTGTAAAGTTT 720
 QY 721 TTCAATGCTACTTGGGCAAGTTGACTGAGTACGAGCAGCATGCTTCTGAGAGCC 780

DB 721 TTCAATGCTACTTGGGCAAGTGTGACTGAGTACGAGCAGCATGCTTCTGAGAGCC 780
 QY 781 GAGGTGACTCTTACATATACAGTGTGCTGCTCCGCGCGTGTGTGGTGCATCACTG 840
 DB 781 GAGGTGACTCTTACATATACAGTGTGCTGCTCCGCGCGTGTGTGTGGTGCATCACTG 840
 QY 841 CTGCTGCGCTGT 900
 DB 841 CTGCTGCGCTGT 900
 QY 901 GAATCTGAAAAGAGAGCAACAAGAACTGAGACAGAAAGGAAATGAAAATC 960
 DB 901 GAATCTGAAAAGAGAGCAACAAGAACTGAGACAGAAAGGAAATGAAAATC 960
 QY 961 CGGCTACAAATTCAGATGAACAAAGACACAGACCGGCTTCTCTCCCTCCAAATCTG 1020
 DB 961 CGGCTACAAATTCAGATGAACAAAGACACAGACCGGCTTCTCTCCCTCCAAATCTG 1020
 QY 1021 TGAATCCAGTGAATCCGAAACAAAGAACTGAGTGTGCTCTACACAGCGGCTGA 1080
 DB 1021 TGAATCCAGTGAATCCGAAACAAAGAACTGAGTGTGCTCTACACAGCGGCTGA 1080
 QY 1081 TCAAGCTCCAGCCAGGCGCAGCAAGTATCCACAGGCTCTTTTAATCTGGCCAGTCTGA 1140
 DB 1081 TCAAGCTCCAGCCAGGCGCAGCAAGTATCCACAGGCTCTTTTAATCTGGCCAGTCTGA 1140
 QY 1141 GAAGTCTAATTAATCACTGATATAG 1168
 DB 1141 GAAGTCTAATTAATCACTGATATAG 1168

RESULT 2
 AX380396 1175 bp DNA linear PAT 18-MAR-2002
 LOCUS AX380396
 DEFINITION Sequence 1 from Patent WO0200710.
 ACCESSION AX380396
 VERSION AX380396.1 GI:19575326
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Weicher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
 B7-like molecules and uses thereof
 Patent: WO 0200710-A 1 03-JAN-2002;
 JOURNAL Amgen, Inc. (US)

FEATURES
 Location/Qualifiers
 1.1175
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 27.1175
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28216.1"
 /db_xref="GI:19575327"
 /translation="MGLVIFLHSGSGSENYIEGPONAVLKGSOARFNCYVSGMKLI
 MMLASDMVLSVREMEPIITNDRETSQYDGGNTSEMIITHNVEPSDGINRC
 SYRVSPEPSDLOSASVILATPOSNGTLCAVATKSKARSATVNLVIRCPDPTDGG
 GINIPGVLSLPSLGEFLPTWKGVLGAGMLLPTCTLTIRCCCRRCGCCNCC
 RCGCRRCRRKRGFRIOFOKSKSEKTKNETETESGNGSYNSDEKTDITSLP
 ESSDPEQNSCCGPHQADRPAPRPSHPQASFNILASPERVNTIV"

BASE COUNT 295 a 307 c 286 g 287 t
 ORIGIN

Query Match 96.1%; Score 1122.4; DB 6; Length 1175;
 Best Local Similarity 99.5%; Pred. No. 2.2e-259;
 Matches 1126; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 37 CCCACCGGTTCTGGGTGTGAATGAAGTATGAAGGCCCCCAAAATGCAAGAGTCT 96
 || | |||||

Db 44 CCTCCACGGTTCTGGGCTGGGTAATGAATCATAGAAAGGCCCCAGAAATGCACAGTCTT 103
 QY 97 GAAGGGCTCCAGAGCTCGCTTCAACTGCACCGTCTCCAGGGGCTGGAAGCTCATCATGTG 156
 Db 104 GAAGGGCTCCAGAGCTCGCTTCAACTGCACCGTCTCCAGGGGCTGGAAGCTCATCATGTG 163
 QY 157 GGCTCTGAGTACATGAGTGTCTTAAGGTGACAGGCGCATGAGGCCATCATCACAATGA 216
 Db 164 GGCTCTGAGTACATGAGTGTCTTAAGGTGACAGGCGCATGAGGCCATCATCACAATGA 223
 QY 217 CGGCTTCACTCTCAGAGGTACAGGAGGCGGGAATTCACCTGGAGATGATATGCA 276
 Db 224 CGGCTTCACTCTCAGAGGTACAGGAGGCGGGAATTCACCTGGAGATGATATGCA 283
 QY 277 CAATGTGAGCCCAAGATTTGGGGGAAATCATGATCAGCTCCAGAACATGCGCTGCA 336
 Db 284 CAATGTGAGCCCAAGATTTGGGGGAAATCATGATCAGCTCCAGAACATGCGCTGCA 343
 QY 337 TGGATCTGCTTACCTTACCGTCCAAATTAAGGAGAGCTGTTCAATCCAGTGTATCT 396
 Db 344 TGGATCTGCTTACCTTACCGTCCAAATTAAGGAGAGCTGTTCAATCCAGTGTATCT 403
 QY 397 TGTAGTGTCTGAGAAATGAACCTTGTGAAGTACTTGTACTACCTCAGCTGAGCTGCT 456
 Db 404 TGTAGTGTCTGAGAAATGAACCTTGTGAAGTACTTGTACTACCTCAGCTGAGCTGCT 463
 QY 457 CCGGATATTTCCGTCGAGAGCTCGTCTCCTGTCAGAGCATTCAGCTATTTTGTTC 516
 Db 464 CCGGATATTTCCGTCGAGAGCTCGTCTCCTGTCAGAGCATTCAGCTATTTTGTTC 523
 QY 517 GGAGCCCAAGGACCTTAAAGTGCAGTACAGTACCTGGCTGAGCCCAAGAGCAATG 576
 Db 524 GGAGCCCAAGGACCTTAAAGTGCAGTACAGTACCTGGCTGAGCCCAAGAGCAATG 583
 QY 577 GACTTTGACTTGGCTGCTACTGGAAGAGCTGAAAGCCCGCAAGTCTCACTGATAA 636
 Db 584 GACTTTGACTTGGCTGCTACTGGAAGAGCTGAAAGCCCGCAAGTCTCACTGATAA 643
 QY 637 TCTCAGTGTGATGCTGCTGCTCCCAAGACACTGGAAGTGTATTAATTCAGGTAT 696
 Db 644 TCTCAGTGTGATGCTGCTGCTCCCAAGACACTGGAAGTGTATTAATTCAGGTAT 703
 QY 697 ATCAAGTTACCGAGTTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 756
 Db 704 ATCAAGTTACCGAGTTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 763
 QY 757 AGCAGGACCAATGCTTCTGAGCGGAGCTGACTTCAATPACGCTGCTGCTGCG 816
 Db 764 AGCAGGACCAATGCTTCTGAGCGGAGCTGACTTCAATPACGCTGCTGCTGCG 823
 QY 817 CCGTGTGTGTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
 Db 824 CCGTGTGTGTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
 QY 877 AGGATTTCTATTCATTTCAAAAGAAATCTGAAAAGAGAACAAACAAAGAACTGA 936
 Db 884 AGGATTTCTATTCATTTCAAAAGAAATCTGAAAAGAGAACAAACAAAGAACTGA 943
 QY 937 GACAGAAATGGAATGAATAATCCGGCTCAATTCAGATGAAAGAAAGAACAGACAC 996
 Db 944 GACAGAAATGGAATGAATAATCCGGCTCAATTCAGATGAAAGAAAGAACAGACAC 1003
 QY 997 CGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
 Db 1004 CGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063
 QY 1057 TGGCCCTCTCAGAGGCGGCTGATCAGCTCAGCCAGGAGCAAGTCAATCCAGAG 1116
 Db 1064 TGGCCCTCTCAGAGGCGGCTGATCAGCTCAGCCAGGAGCAAGTCAATCCAGAG 1123
 QY 1117 TTTCTTTAATCTGCGCAGTCTGAGAAAGTCAATATACACTGTAATAG 1168
 Db 1124 TTTCTTTAATCTGCGCAGTCTGAGAAAGTCAATATACACTGTAATAG 1175

RESULT 3
 AK092516 2051 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ35197 f1s, clone PLACE6017788, highly similar
 DEFINITION to IGSE5.
 ACCESSION AK092516
 VERSION AK092516.1 GI:21751130
 KEYWORDS oligo capping, f1s (full insert sequence).
 SOURCE Homo sapiens placenta cDNA to mRNA, clone_11b:PLACE6
 clone:PLACE6017788.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1. Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Magatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Negai, K.
 and Isogai, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2051)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA library
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 Location/Qualifiers
 source
 1..2051
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE6017788"
 /tissue_type="Placenta"
 /clone_lib="PLACE6"
 /note="cloning vector: pME18SFU3"
 BASE COUNT 580 a 463 c 474 g 534 t
 ORIGIN
 Query Match 95.9%; Score 1119.6; DB 9; Length 2051;
 Best Local Similarity 99.6%; Pred. No. 1.1e-258;
 Matches 1122; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 43 CGGTTCTGGGCTGGTATGAAGTCAATGAAGGCCCAAAATGCAAGAGTCTCTGAAGG 102
 Db 202 CGGTTCTGGGCTGGTATGAAGTCAATGAAGGCCCAAAATGCAAGAGTCTCTGAAGG 261
 QY 103 CTCCAGGCTCGCTTCAATGACACCGTCTCCAGGGCTGGAAGTCAATGATGGGCTCT 162
 Db 262 CTCCAGGCTCGCTTCAATGACACCGTCTCCAGGGCTGGAAGTCAATGATGGGCTCT 321
 QY 163 CAGTACATGTTGGTGTAAAGCTCAGAGCCCATGAGCCCATCATCAATGACAGGCTT 222
 Db 322 CAGTACATGTTGGTGTAAAGCTCAGAGCCCATGAGCCCATCATCAATGACAGGCTT 381
 QY 223 CACCTCTCAGAGTACAGCAGGCGGGAATCTTCACTCGGAGATGATCATCCCAATGT 282
 Db 382 CACCTCTCAGAGTACAGCAGGCGGGAATCTTCACTCGGAGATGATCATCCCAATGT 441
 QY 283 GGAGCCCAAGTATTTGGGGAACATCAGATGAGCTCCAGAAAGTCCCTGATGATC 342
 Db 442 GGAGCCCAAGTATTTGGGGAACATCAGATGAGCTCCAGAAAGTCCCTGATGATC 501

yY
 1 ATGGTGGCAGGAGCCATGGAAAAATAGAGACCAACCACCGGTTGTGGTCTGGTAATGAAGTC 60
 Db
 yY
 68 ATGAAAGGCCCCCAAAATGCAGAGTCCGTGAAAGGGCTCCAGAGTCGCTTCAACTGCACC 127

[illegible]

RESULT 6
AX380404
LOCUS AX380404 1195 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0200710.
ACCESSION AX380404
VERSION AX380404.1 GI:19575334
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 9 03-JAN-2002;
Amgen, Inc. (US)

FEATURES
source location/Qualifiers
1..1195
/organism="Mus musculus"
/db_xref="taxon:10090"
53..1165
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28220.1"
/db_xref="GI:19575335"
/translation="MEGSRDYLAVLVLAQDLSGSSYQIIEGPNVTVLKDSEAH
DSGVVCSLQNSHSGSALFSLVQVMTLNINPNIATYEGPCVTCYAGCMSTLPI
SWLEIYVPSHSYNSFLFENFMVLSVLDLPLGNTLCVAFELDLQASKEITVNL
TYVQPPDSIGEGSPALPTMAIILAVAFSLILILVILITICCCASRREESTY
QNEIKRSAMRNKADPETKLKSGKENYSSDPAKKAQASLIPKSAEVSLEPKRSS
SLPYELNKHQGPATHRVSEFASIPKRVNTLV"

BASE COUNT 313 a 304 c 291 g 287 t

Query Match
Best Local Similarity 32.4%; Score 378.4; DB 6; Length 1195;
Matches 701; Conservative 0; Mismatches 356; Indels 69; Gaps 4;

42 CCGCTCTGGTCTGTGTAATGATCATAGAGAGCCCAAAATGCAAGAGTCTCTAAG 101
108 CAGCTTCGGATCCAGTATCATGATCATAGAGTCTCTAAGATGATCATAGAG 167
102 GCTCCAGGCTGCTTCAACGACGCTCTCCAGGCTGGAAGCTCATCATGTGGGCTC 161
168 ACTCAGAGGCTCATTCACTGACGCTGACGCTGGAAGCTTCTCATGTGGACTC 227
162 TCAGTCAATGTGTGTCTAGCGCTCAGGCCCATGAGCCATCATCATGACCATG 221
228 TTAACCAATGTGTGTCTAGCTGCTCAGCCAGGACCATCATCATCAACACCGCT 287
222 TCACCTCAGAGGTAGCAGCAGGCGGAGACTTCACTGAGAGATGATCATCCCAATG 281
288 TCACCTATGCCAGTTCACACAGACGATGATCTCATCTCGGAGTTCATTCATCATGATG 347
282 TGGAGCCCATGATGCGGAGACATCATGAGCTCCAGAACAGTCCCTGATGATG 341
348 TGCAGCCCAATGTGTGTCTAGCGCTCAGGCCCATGAGCCATCATCATCAACACCGCT 407
342 CTGCTTACCTACCGTCCAGATTTATGAGAGAGCTTTATTTCCAGTGTAAATTTAG 401
408 CTGCTTACCTACCGTCCAGATTTATGAGAGAGCTTTATTTCCAGTGTAAATTTAG 467
402 TGGCTAGATGACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 461
468 TCACTAGAGGTGACCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 527
462 AATATTCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 521
528 AATATTCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587

QY 522 CCAGGACCTTCAAGTGCAGTGCATCTGCTGATGCCCAAGACAGCAATGAGACTT 581
Db 588 CGGCACTTTATAGAGGCTCTTGAGTCTCTGAGCTCAGACCTCAGACAGTGCAGGAGACT 647
QY 582 TGACTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 641
Db 648 TGACTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
QY 642 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701
Db 708 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
QY 702 GTTATCCGAGTTCATGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 761
Db 745 -----AGCCGAGCAGCTGCGGACCTGCGGACCTATCTCTGCTGCTGCTG 791
QY 762 GCACATGCTTTCAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821
Db 792 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
QY 822 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
Db 844 -----TTGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
QY 882 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 941
Db 899 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
QY 942 AAGTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1001
Db 945 AAGTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1004
QY 1002 CTCTCCCTCCCAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTG 1061
Db 1005 CTCTCCCTCCCAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTG 1058
QY 1062 CTCTCCCTCCCAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTG 1121
Db 1059 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
QY 1122 TTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1167
Db 1119 TTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1164

RESULT 7
AF121782
LOCUS AF121782 142742 bp DNA linear PRI 02-FEB-1999
DEFINITION Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.
ACCESSION AF121782
VERSION AF121782.1 GI:4210991
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 142742)
AUTHORS Taudien, S., Dagand, E., Hildmann, T., Nordsiek, G., Drescher, B.,
Schaltevoy, R., Weber, J., Schilling, M., Menzel, U., Yaspo, M.L. and
Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1999) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES
source location/Qualifiers
1..142742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 206A10"
3..118
/evidence="not_experimental"
/rpe_family="MIR"


```

repeat_region complement(140..415)
/evidence-not_experimental
/rpt_family="MLT2E"
complement(668..881)
/evidence-not_experimental
/rpt_family="MER38A"
complement(1078..1162)
/exon /note="GRAIL, score = 86.000%, comment = excellent shadow"
/evidence-not_experimental
1114..1403
repeat_region /evidence-not_experimental
/rpt_family="MLT1I"
repeat_region /evidence-not_experimental
1481..1603
/evidence-not_experimental
/rpt_family="L2"
1576..1710
/exon /note="MZE, score = 66.1%"
/evidence-not_experimental
1576..1710
/exon /note="GRAIL, score = 99.000%, comment = excellent"
/evidence-not_experimental
complement(1693..1741)
repeat_region /evidence-not_experimental
/rpt_family="MR"
complement(1780..2326)
repeat_region /evidence-not_experimental
/rpt_family="MLT1E"
/evidence-not_experimental
complement(1834..1955)
/exon /note="GRAIL, score = 50.000%, comment = good shadow"
/evidence-not_experimental
2629..2831
repeat_region /evidence-not_experimental
/rpt_family="ALUSX"
3012..3122
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"
3123..3478
repeat_region /evidence-not_experimental
/rpt_family="THE1B"
3479..3618
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"
complement(3502..3540)
/exon /note="GRAIL, score = 65.000%, comment = good shadow"
/evidence-not_experimental
4674..4781
/exon /note="GRAIL, score = 42.000%, comment = marginal"
/evidence-not_experimental
4712..4749
/exon /note="Xpound exon prediction, score = 70% (0%)"
/evidence-not_experimental
5385..6104
repeat_region /evidence-not_experimental
/rpt_family="HERV1"
6682..6751
/exon /note="MZE, score = 95.1%"
/evidence-not_experimental
6746..6751
/exon /note="Genscan, score = 1.83%, comment = Initial-exon 6 bp
frame: 1 phase: 0"
/evidence-not_experimental
6901..7339
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
7347..7457
repeat_region /evidence-not_experimental
/rpt_family="LRR16C"
complement(7561..7696)
repeat_region /evidence-not_experimental
/rpt_family="MR"
complement(7819..8220)
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"

exon 8097..8119
/note="GRAIL, score = 45.000%, comment = marginal"
/evidence-not_experimental
complement(8738..9232)
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
complement(9273..9420)
/exon /note="GRAIL, score = 71.000%, comment = good"
/evidence-not_experimental
9290..9371
/exon /note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence-not_experimental
9966..10041
repeat_region /evidence-not_experimental
/note="homology = 100.00%, score = 38, counts = 2"
/evidence-not_experimental
/rpt_type=tandem
/rpt_unit=ctcttcacactctctgtgagagcggttatcgc
10244..10664
repeat_region /evidence-not_experimental
/rpt_family="MLT1E"
10665..11028
repeat_region /evidence-not_experimental
/rpt_family="MLT1F"
11318..11417
repeat_region /evidence-not_experimental
/rpt_family="L1MC/D"
11503..11818
repeat_region /evidence-not_experimental
/rpt_family="ALUSG"
complement(11525..11777)
/exon /note="MZE, score = 79.4%"
/evidence-not_experimental
complement(12346..12827)
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
12902..13180
repeat_region /evidence-not_experimental
/rpt_family="ALUSG"
complement(13057..13161)
/exon /note="MZE, score = 76.1%"
/evidence-not_experimental
complement(13226..13364)
repeat_region /evidence-not_experimental
/rpt_family="L1MC2"
complement(13613..13908)
repeat_region /evidence-not_experimental
/rpt_family="ALUSX"
complement(14317..14614)
repeat_region /evidence-not_experimental
/rpt_family="LRR37A"
complement(14962..15036)
/exon /note="MZE, score = 93%"
/evidence-not_experimental
15255..15299
repeat_region /evidence-not_experimental
/rpt_family="L1PA6"
complement(15469..15497)
/exon /note="Xpound exon prediction, score = 62% (0%)"
/evidence-not_experimental
16045..16768
repeat_region /evidence-not_experimental
/rpt_family="MER21A"
complement(16555..16655)
/exon /note="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
16823..16850
repeat_region /evidence-not_experimental
/rpt_family="MER34"
complement(16851..17208)
repeat_region /evidence-not_experimental
/rpt_family="THE1A"
17209..17632
repeat_region /evidence-not_experimental

```


	/note="L2"	/rpl_family="LINE/L2"	/rpl_type=DISPERSED	complement(10078. .10429)	
repeat_region	/note="THEIB"	/rpl_family="LTR/MALR"	/rpl_type=DISPERSED	complement(10430. .10651)	
repeat_region	/note="L2"	/rpl_family="LINE/L2"	/rpl_type=DISPERSED	complement(11002. .11282)	
repeat_region	/note="(TTTA)n"	/rpl_family="Simple_repeat"	/rpl_type=DISPERSED	10980. .11000	
repeat_region	/note="AluSp"	/rpl_family="SINE/Alu"	/rpl_type=DISPERSED	complement(11315. .11556)	
repeat_region	/note="AluJo"	/rpl_family="SINE/Alu"	/rpl_type=DISPERSED	11702. .11753	
repeat_region	/note="MIR"	/rpl_family="SINE/MIR"	/rpl_type=DISPERSED	12649. .12678	
exon	/gene="SH3BGR"	/number=5	complement(13115. .13425)		
repeat_region	/note="Alusx"	/rpl_family="SINE/Alu"	/rpl_type=DISPERSED	complement(13598. .13892)	
repeat_region	/note="Aluy"	/rpl_family="SINE/Alu"	/rpl_type=DISPERSED	14013. .14131	
repeat_region	/note="(CA)n"	/rpl_family="Simple_repeat"			
	28.0%; Score 326.6; DB 9;	Length 340000;			
Query Match	Best Local Similarity 97.4%; Pred. No. 9.8e-68;	Mismatches 9;	Gaps 0		
Matches 332; Conservative					
QY 27	AAATAGAGACCCACCGGCTTCGGTGGGTGAATGAAGTCATAGAACGCCCAAATG	86			
Db 268437	AAATTGCCTACCTTCAGAGTCTGGGTCTGGTAATGAAGTCATAGAACGCCCAAATG	268495			
QY 87	CAGAAGCTCGAAGGGCTCCAGAGCTCGCTCAACTGCACCGCTCCACAGGGCTGGAAAGC	146			
Db 268497	CMAAGCTCGAAGGGCTCCAGAGCTCGCTCAACTGCACCGCTCCACAGGGCTGGAAAGC	268556			
QY 147	TGATCATGTGGGCTCTCAGTAGAGATGATGTTGGTGTAAAGCTCAGGGCCATGAGCCATCA	206			
Db 268557	TGATCATGTGGGCTCTCAGTAGAGATGATGTTGGTGTAAAGCTCAGGGCCATGAGCCATCA	268616			
QY 207	TCACCAGTGACCGGCTTACCTCTCAGAGTAGACACACAGGGCGGGAACCTCACCTGGGAGA	266			
Db 268617	TCACCAGTGACCGGCTTACCTCTCAGAGTAGACACACAGGGCGGGAACCTCACCTGGGAGA	268676			
QY 267	TGATCATCCACAATGTGAGAGCCCAAGTATTCGGGGAGACATCAATGCAGGCTTCAGAGACA	326			
Db 268677	TGATCATCCACAATGTGAGAGCCCAAGTATTCGGGGAGACATCAATGTGAGGCTTCAGAGACA	268736			
QY 327	GTCGCTGCATGATCTGCTTACCTTACCTTACCGTCAAGTTATG	367			
Db 268737	GTCGCTGCATGATCTGCTTACCTTACCTTACCGTCAAGTTATG	268777			
RESULT 9					
ACO20851					

LOCUS AC020851 196900 bp DNA linear HTG 15-JUL-2000
 DEFINITION Mus musculus clone RP21-467112, WORKING DRAFT SEQUENCE, 33
 AC020851
 AC020851.2 GI:9211212
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 196900)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Mouse
 JOURNAL Unpublished
 2 (bases 1 to 196900)
 DIRECT SUBMISSION
 TITLE DOE Joint Genome Institute.
 JOURNAL Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jul 15, 2000 this sequence version replaced gi:6686457.
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 1437240
 Center clone name: RP21-467112

Summary Statistics
 Consensus quality: 164618 bases at least Q40
 Consensus quality: 180887 bases at least Q30
 Consensus quality: 183777 bases at least Q20
 Estimated insert size: 147000; pulse field gel estimation
 Quality coverage: 5.67 in Q20 bases; sum-of-coverage estimation
 Quality coverage: 4.3 in Q20 bases; pulse field gel estimation
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 1541: contig of 1541 bp in length
 1542 1641: gap of unknown length
 1642 3492: contig of 1851 bp in length
 3493 3592: gap of unknown length
 3593 4738: contig of 1146 bp in length
 4739 4839: gap of unknown length
 4839 7249: gap of unknown length
 7249 7348: gap of unknown length
 7348 9178: contig of 1830 bp in length
 9179 9279: gap of unknown length
 9279 11493: contig of 2215 bp in length
 11494 11593: gap of unknown length
 11594 13607: contig of 2014 bp in length
 13608 13707: gap of unknown length
 13708 15946: contig of 2239 bp in length
 15947 16046: gap of unknown length
 16046 18392: contig of 2346 bp in length
 18393 18492: gap of unknown length
 18493 21231: contig of 2739 bp in length
 21232 21331: gap of unknown length
 21332 24662: contig of 3331 bp in length
 24663 24763: gap of unknown length
 24763 27140: contig of 2378 bp in length
 27141 27240: gap of unknown length
 27240 29950: contig of 2710 bp in length
 29951 30050: gap of unknown length
 30051 32245: contig of 2195 bp in length
 32246 32345: gap of unknown length

32346 35688: contig of 3343 bp in length
 35689 35788: gap of unknown length
 35789 40387: contig of 4599 bp in length
 40388 40487: gap of unknown length
 40488 43580: contig of 3093 bp in length
 43581 43680: gap of unknown length
 43681 46437: contig of 2757 bp in length
 46438 46537: gap of unknown length
 46538 50184: contig of 3647 bp in length
 50185 50284: gap of unknown length
 50285 54572: contig of 4288 bp in length
 54573 54672: gap of unknown length
 54673 60327: contig of 5555 bp in length
 60328 60328: gap of unknown length
 60328 68024: contig of 7697 bp in length
 68025 68124: gap of unknown length
 68125 78610: contig of 10486 bp in length
 78611 78710: gap of unknown length
 78711 88151: contig of 9441 bp in length
 88152 88251: gap of unknown length
 88252 96257: contig of 8006 bp in length
 96258 96358: gap of unknown length
 96359 106327: contig of 9870 bp in length
 106328 106327: gap of unknown length
 106327 115795: contig of 9468 bp in length
 115796 115795: gap of unknown length
 115796 126363: contig of 10468 bp in length
 126364 126463: gap of unknown length
 126464 137189: contig of 10726 bp in length
 137190 137289: gap of unknown length
 137290 146378: contig of 9089 bp in length
 146379 146478: gap of unknown length
 146479 161686: contig of 15208 bp in length
 161687 161786: gap of unknown length
 161787 176149: contig of 14363 bp in length
 176150 176249: gap of unknown length
 176250 196900: contig of 20651 bp in length.

FEATURES

1..196900
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP21-467112"
 /clone.lib="RP21 mouse PAC library 21"

BASE COUNT 51456 a 43823 c 44608 g 53666 t 3347 others

Query Match

Best Local Similarity 72.3%; Score 182.4; DB 2; Length 196900;
 Pred. No. 4.8e-33; Matches 237; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 37 CCCACCGGTTCTGGTGTGTAATGAAGTCAATGAGGCCGCCCAAAATGACAGTCTCT 96
 DB 67439 CCTTCTAGCTTCGGATTCAGTATTCAGATCATATGAAAGTCTCAGATTAACAGTCT 67498
 QY 97 GAAGGCTCCCAAGGCTCCCTTCACTGACCGTCTCCAGGCGTGAAGCTCATATG 156
 DB 67499 AAAGAGCTAGAGGCTGCTCACTGACCGTCTGACCGTGAAGCTTCATG 67558
 QY 157 GCGTTCAGTACATGTTGGTGTGTAAGGCTGACGCGCATGAGCCATCATCATCA 216
 DB 67559 GACTTCTAACCAATGTTGGTGTGTAAGGCTGACGCGCATGAGCCATCATCA 67618
 QY 217 CCGTTCAGCTTCAGAGGTCACACGAGGCGGGAATTCCTCCTGAGATGATCA 276
 DB 67619 CCGTTCAGCTTCAGAGGTCACACGAGGCGGGAATTCCTCCTGAGATGATCA 67678
 QY 277 CAATGTGAGCCAGTATGTTGGGGAACATCAATGATGATGATGATGATGATGAT 336
 DB 67679 TGATGTGAGCCAGTATGTTGGGGAACATCAATGATGATGATGATGATGATGAT 67738
 QY 337 TGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364
 DB 67739 TGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 10	AX380408	754 bp	DNA	linear	PAT 18-MAR-2002
LOCUS	AX380408				
DEFINITION	Sequence 13 from Patent WO0200710.				
ACCESSION	AX380408				
VERSION	AX380408.1	GI:19575338			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Weilcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.				
TITLE	B7-like molecules and uses thereof				
JOURNAL	Patent: WO 0200710-A 13 03-JUN-2002;				
FEATURES	Amgen, Inc. (US)				
source	Location/Qualifiers				
	1..754				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	53..724				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD28222.1"				
	/db_xref="GI:19575338"				
	/translation="MEGSRDVLAVLVILVLAQLRASSGYIIEGPNVYLKRSFAH				
	NCVTWGLKLMWTLNQWVYLSLTGPIITNNKFTYASTNSDSFISELLIHDPYS				
	DSGAVQCSLQNSHGSGAFSLVQESTYSEIIRKSNARTKAPETPLKSGKNTGTS				
	SDENAAQGTSLSPKSAVSLPEKRSSLPYQELNKHQPGPATHPVSPDIASPKVR				
	NVTLY"				
BASE COUNT	220 a 191 c 175 g 168 t				
ORIGIN					
Query Match	15.6%; Score 181.8; DB 6; Length 754;				
Best Local Similarity	72.9%; Pred. No. 5, 1e-33;				
Matches 234; Conservative	0; Mismatches 87; Indels 0; Gaps 0;				
QY	42 CCGGTTCTGGCTGTGTAATGAGTCATATGAAAGGCCCCCAAAATGCAAGTCTGAAGG	101			
DB	108 CAGGTTCCGATTCAGATTACATGATCATGAAAGGTCCTTAGAATGTAAACAGTCCAAAGG	167			
QY	102 GCTCCAGGCTCGCTTCATCTCACTGACACGCTCCCGAGGCTGGAAGCTCATCATGTGGCTC	161			
DB	168 ACTAGAGGCTCACTTAATGACACGCTGATCAGCGCTGGAAGCTTTCATAGTGGATC	227			
QY	162 TCAGTGAATGTGTGCTTAAGCTCAAGGCCCATGGAAGCCCATCATCACCATAAGACGCT	221			
DB	228 TTAAACCAAAATGTGTGCTAGTGTCCACCAACCAAGGACCATCATCATCCACCAACCGCT	287			
QY	222 TCACCTTCAGAGGTAGACACAGGCGGAGAACTTCACCTTGAGATGATCATCCACATG	281			
DB	288 TCACCTATGCCAGTTACACAGCACTGACAGCTTCATCTCGAGTTGATCATCATCATGAG	347			
QY	282 TGGAGCCCAATGATTCGGGGAACATCAGATGCAAGCCCTCCGAAACAGTCCGCTCATGAT	341			
DB	348 TGCACCCCAAGTACCGGAGTCCGTGCATGACAGCTTCGAAACAGCCATGGGTTTGAT	407			
QY	342 CTGCTTACCTTACCGTCCAG 362				
DB	408 CTGCTTCTCTCTCAGTGCAG 428				
RESULT 11					
LOCUS	AX380406	895 bp	DNA	linear	PAT 18-MAR-2002
DEFINITION	Sequence 11 from Patent WO0200710.				
ACCESSION	AX380406				
VERSION	AX380406.1	GI:19575336			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 Melcher, A.A., Sarmiento, U.M., Schultze, H.J. and Chute, H.T.			
TITLE	B7-like molecules and uses thereof			
JOURNAL	Patent: WO 0200710-A 11 03-JAN-2002;			
Amgen, Inc. (US)	Location/Qualifiers			
source	1. 895			
CDS	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	53. . 865			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAD28221.1"			
	/db_xref="GI:19575337"			
	/translation="MEGSRMDYTAIVVILNQLTAGSSVYQIEGPONTVLEKSEAHF			
	NCTVTHGKLLMTNOMVSLTGTGPTLNTNRFVASTNDSFISPLIHVDPS			
	DSGTVOCILSHQFSGFAFLSVSDSIGSESPALPMAIILIAVATSLIILIVLI			
	FCCCAARREKEESTYONEIKRSANMNTNPADEFTLKSGKEVYCGSSPEAKAQTAS			
	LPFSAEVSLEPKRSSSLPYOELNKHQPGPATPRVSPFSDIASFOKVRNTLV"			
BASE COUNT	250 a	223 c	212 g	210 t
ORIGIN				
Query Match	15.6%	Score 181.8;	DB 6;	Length 895;
Best Local Similarity	72.9%;	Prod. No. 5.2e-33;		
Matches 234;	Conservative	0;	Mismatches 87;	Indels 0;
				Gaps 0;
Qy	42	CGGGTCTGGTCTGGTATGATGAGTCAAGAGGCCCCCAATGCAAGACTCTGAAGG	101	
Db	108	CAGCTTCGGAGTCAAGTTATGAGATCATGAGAGGCTCTCAAAATGTAACATCTCAAGG	167	
Qy	102	GCTCCAGGCTCGCTTCAACTGACACCGCTCCCGAGGGTGGAGAGTCAATCATGTGGGCTC	161	
Db	168	ACCTGAGGGCTCACTTCAACTGACCGTACCGTACCTGAGAGTCTGAGAGCTTCTCATGTGGACTC	227	
Qy	162	TCAGTGCATGATGGTGTCTAAGGCTCAGGCCCATGAGGCCATCATCTACCAATGACCCCT	221	
Db	228	TTAACCAAAAGTGGTGGTGGAGTCTCACCACCAAGACCCATCTCATCAACCAACCCCT	287	
Qy	222	TCACCTCTCAGAGGTACGACAGAGGGGGAACCTTCACTCGAGATGATCATCCCAATG	281	
Db	288	TCACCTATGGCACTTTCACACGACATGACAGCTTCACTCTCGAGTTCATCATTCATGATG	347	
Qy	282	TGAGGCCAGTGTGGGGAGACATGATGACAGCTTCAGAACAGTCCCTCGCATGAT	341	
Db	348	TGCAGCCCACTGACTCGGGATCGGTGGAATGACAGCTCGAGAACACCATGGGTTGGAT	407	
Qy	342	CTGCTTACCTTACCCTCCAGG	362	
Db	408	CTGCCTTCCTCTCAGTCAAG	428	
RESULT 12				
BC004806	1556 bp	mrna	linear	ROD 07-AUG-2002
LOCUS				
DEFINITION	Mus musculus, RIKEN cDNA 2010003D20 gene, clone MGC:7960			
IMAGE:3584645,	mrna, complete cds.			
ACCESSION	BC004806			
VERSION	BC004806.1 GI:13435932			
KEYWORDS	MGC.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 Strausberg, R.			
AUTHORS	Direct Submission			
TITLE	Submitted (21-MAR-2001) National Institutes of Health, Mammalian			
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
USA				
NIH-MGC Project URL:	http://mgc.nci.nih.gov			


```

source
1. 175861
   /organism="Mus musculus"
   /db_xref="taxon:10090"
   /clone="Rnp3-147E11"
   /clone_id="RPCI-23 Female Mouse BAC"
   1. 28835
   /note="assembly_fragment"
   /note="assembly_fragment"
   clone_end:Sp6
   vector_side:left"
   28936..29562
   /note="assembly_fragment"
   29663..30328
   /note="assembly_fragment"
   30429..32005
   /note="assembly_fragment"
   32106..33128
   /note="assembly_fragment"
   33229..34512
   /note="assembly_fragment"
   34613..36079
   /note="assembly_fragment"
   36180..37483
   /note="assembly_fragment"
   37584..39166
   /note="assembly_fragment"
   39267..41065
   /note="assembly_fragment"
   41166..43203
   /note="assembly_fragment"
   43304..47489
   /note="assembly_fragment"
   47590..50604
   /note="assembly_fragment"
   50705..56370
   /note="assembly_fragment"
   56471..64079
   /note="assembly_fragment"
   64180..72619
   /note="assembly_fragment"
   72720..81566
   /note="assembly_fragment"
   81667..96774
   /note="assembly_fragment"
   96875..119647
   /note="assembly_fragment"
   119748..147772
   /note="assembly_fragment"
   147873..174557
   /note="assembly_fragment"
   174658..175861
   /note="assembly_fragment"
   clone_end:r7
   vector_side:right"

BASE COUNT      48903 a 39540 c 39029 g 46286 t 2103 others
ORIGIN
Query Match      14.7% ; Score 171.4 ; DB 2 ; Length 175861 ;
Best Local Similarity 72.0% ; Pred. No. 2.le-30 ;
Matches 237 ; Conservative 0 ; Mismatches 91 ; Indels 1 ; Gaps 1 ;

1
37 CCCACCGCGTTCTGGGTCTGCTAATGAGAGTCAATGAGCCGCCCAAAATGCAAGAGTCT 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120335 CATTCTAGCTTCGGAGTTCAGTTATCAGTCTTAAGAGTCTCTAGATGTACAGTCCT 120476

97 GAAGGGTCCAGCGCTCGCTCAACTGACCGT-CTCCAGGGGTGAAGCTCATCTATGT 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120475 AAAGGACTCAGAGGCTCATTCACTGACGACCGTGGGTACGAGTGAAGCTTCTCATGT 120416

156 GGGCTCTCACTGACATGCTAGTCTCTAAGCGCTAGGCCCATGAGCCCATCATCCCAATG 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120415 GGACTTTAACCAAAATGGTGCTGTAAGTCTACACCAACCAAGAACCCATCTACACAA 120356

216 ACCGCTTACCTCTCAGAGGTGACGACCAAGGGGGGAAGTTCACCTCGAGATGATCATCC 275

```



```

repeat_region 33468..33678
/evidence-not_experimental
/rpt_family="THELB"
34563..34620
/notes="homology = 93.10%, score = 21, counts = 2"
/evidence-not_experimental
/rpt_type="candem"
/rpt_unit="cttcgaagaagacgacttaggca
repeat_region complement(35785..35824)
/evidence-not_experimental
/rpt_family="MIT2CB"
complement(35924..36345)
/evidence-not_experimental
/rpt_family="MIT2CB"
complement(36403..36475)
/evidence-not_experimental
/rpt_family="MIT2CB"
complement(36717..36828)
/evidence-not_experimental
/rpt_family="MER67C"
complement(37064..37364)
/evidence-not_experimental
/rpt_family="AluY"
37440..37501
/notes="homology = 98.40%, score = 29, counts = 2"
/evidence-not_experimental
/rpt_type="candem"
/rpt_unit="aaacactcacttgcttcaaacctgc
repeat_region 37518..38158
/evidence-not_experimental
/rpt_family="L2"
38477..39376
/evidence-not_experimental
/rpt_family="L2"
39552..39861
/evidence-not_experimental
/rpt_family="MITID"
39884..40122
/evidence-not_experimental
/rpt_family="MITID"
complement(40243..40542)
/evidence-not_experimental
/rpt_family="AluX"
40343..41180
/evidence-not_experimental
/rpt_family="L2"
complement(42066..42114)
/evidence-not_experimental
/rpt_family="MIR"
42138..42317
/evidence-not_experimental
/rpt_family="L2"
complement(42332..42710)
/evidence-not_experimental
/rpt_family="L1P81"
complement(42758..46751)
/evidence-not_experimental
/rpt_family="L1P81"
complement(46792..47084)
/evidence-not_experimental
/rpt_family="L1MA"
47189..47992
/evidence-not_experimental
/rpt_family="L2"
complement(48002..48132)
/evidence-not_experimental
/rpt_family="MIR"
48905..49310

```

Query Match 8.6%; Score 100; DB 9; Length 19965;
 Best Local Similarity 100.0%; Pred. No. 3.3e-13;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1069 CCAGCGGGCTGATCAGCTCCACCCAGCCAGCAGATCATCCAGGCTCTTTAACT 1128
|||||
Db 17555 CCAGCGGGCTGATCAGCTCCACCCAGCCAGCAGATCATCCAGGCTCTTTAACT 17614
|||||
QY 1129 GGCCAGTCTGAGAGGTCACTATACACTGTGTATAG 1168
|||||
Db 17615 GGCCAGTCTGAGAGGTCACTATACACTGTGTATAG 17654
|||||

RESULT 15
AC020851/c 196900 bp DNA linear HTG 15-JUL-2000
LOCUS Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33
DEFINITION unordred pieces.
ACCESSION AC020851
VERSION AC020851.2 GI:9211212
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE DOE Joint Genome Institute.
TITLE DOE Joint Genome Institute.
AUTHORS 2 (bases 1 to 196900)
JOURNAL Unpublished
REFERENCE DOE Joint Genome Institute.
TITLE DOE Joint Genome Institute.
AUTHORS 2 (bases 1 to 196900)
JOURNAL Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 15, 2000 this sequence version replaced gi:6686457.
-----Genome Center
Center: Joint Genome Institute
Center code: JGI
Web site: http://www.jgi.doe.gov

-----
Project Information
Center Project Name: 1437240
Center clone name: RPCI-21_467L12
-----
Summary Statistics
Consensus quality: 164618 bases at least Q40
Consensus quality: 180887 bases at least Q30
Consensus quality: 183777 bases at least Q20
Estimated insert size: 147000; pulse field gel estimation
Estimated insert size: 193700; sum-of-ctnigs estimation
Quality coverage: 5.67 in Q20 bases; pulse field gel estimation
Quality coverage: 4.3 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1541: contig of 1541 bp in length
1542 1541: gap of unknown length
1542 3492: contig of 1851 bp in length
1542 3493 3592: gap of unknown length
3493 3592: gap of unknown length
3493 4738: contig of 1146 bp in length
4738 4739 4838: gap of unknown length
4739 4839 7248: contig of 2410 bp in length
4839 7249 7348: gap of unknown length
7249 7349 9178: contig of 1830 bp in length
7349 9179 9278: gap of unknown length
9179 9279 11493: contig of 2215 bp in length
9279 11494 11593: gap of unknown length
11494 11593 13607: contig of 2014 bp in length
11594 13607 13707: gap of unknown length
13607 13707 15946: contig of 2239 bp in length
13708 15946 16046: gap of unknown length
15947 16046 18392: contig of 2346 bp in length
16047 18392: gap of unknown length

```

Search completed: April 29, 2003, 14:30:21
Job time : 4416.94 secs

```
* 18393 18492: gap of unknown length
* 18493 21231: contig of 2739 bp in length
* 21232 21331: gap of unknown length
* 21332 24662: contig of 3331 bp in length
* 24663 27140: contig of 2378 bp in length
* 27141 27240: gap of unknown length
* 27241 29950: contig of 2710 bp in length
* 29951 30050: gap of unknown length
* 30051 32245: contig of 2195 bp in length
* 32246 32343: gap of unknown length
* 32346 35688: contig of 3343 bp in length
* 35689 35788: gap of unknown length
* 35789 40387: contig of 4599 bp in length
* 40388 40487: gap of unknown length
* 40488 43580: contig of 3093 bp in length
* 43581 43680: gap of unknown length
* 43681 46437: contig of 2757 bp in length
* 46438 46537: gap of unknown length
* 46538 50184: contig of 3647 bp in length
* 50185 50284: gap of unknown length
* 50285 54572: contig of 4288 bp in length
* 54573 54672: gap of unknown length
* 54673 60227: contig of 5555 bp in length
* 60228 60327: gap of unknown length
* 60328 68024: contig of 7697 bp in length
* 68025 68124: gap of unknown length
* 68125 78610: contig of 10486 bp in length
* 78611 78710: gap of unknown length
* 78711 88151: contig of 9441 bp in length
* 88152 88251: gap of unknown length
* 88252 96257: contig of 8006 bp in length
* 96258 96357: gap of unknown length
* 96358 106227: contig of 9870 bp in length
* 106228 106327: gap of unknown length
* 106328 115795: contig of 9468 bp in length
* 115796 115895: gap of unknown length
* 115896 126363: contig of 10468 bp in length
* 126364 126463: gap of unknown length
* 126464 137189: contig of 10726 bp in length
* 137190 137289: gap of unknown length
* 137290 146378: contig of 9089 bp in length
* 146379 146478: gap of unknown length
* 146479 161686: contig of 15208 bp in length
* 161687 161786: gap of unknown length
* 161787 176149: contig of 14363 bp in length
* 176150 176249: gap of unknown length
* 176250 196900: contig of 20651 bp in length.
```

FEATURES
source
1. 196900
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP21-467L12"
/clone_lib="RPCR mouse PAC library 21"

BASE COUNT 51456 a 43823 c 44608 g 53666 t 3347 others
ORIGIN

Query Match 4.9%; Score 57.4; DB 2; Length 196900;
Best Local Similarity 66.7%; Pred. No. 0.0058;
Matches 82; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 1045 AACAGTAGCTGCGCCCTCCATCCAGCGGGGCTGATCAAGTCACCGCCAGCGCAGCAG 1104
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 119051 ACACTGACCTTCTACCTTGTATACAGAGCAATATAACATCACGCCGCTCCAGCAGC 118992
QY 1105 TCATCCACAGGCTTCTTTATCTGCGCCAGTCTCTGAGAAAGTGTAGTAATACACTGACT 1164
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 118991 TCATCCACAGGCTTCTTTACATCGCCAGTCTCTCAGAAAGTGTAGTAATGACTTACT 118932
```

QY 1165 ATA 1167
| |
Db 118931 GTA 118929

* . . *

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:24:30 ; Search time 10.2617 seconds
(without alignments)
1543.990 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059
Sequence: 1 MGVIPLHSGSGNEVIEGP.....HPOASFNLASPERKVSNTTVV 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147.5	7.2	404	1 RAGE_HUMAN	Q15109 homo sapien
2	131	6.4	337	1 G55A_CHICK	Q98892 gallus gall
3	129.5	6.3	344	1 MTR1_RAT	Q62718 rattus norv
4	128	6.2	353	1 CEPV_CHICK	Q90773 gallus gall
5	128	6.2	1447	1 DCC_HUMAN	P43146 homo sapien
6	128	6.2	1914	1 KMLS_HUMAN	O15746 homo sapien
7	127	6.2	1447	1 DCC_MOUSE	P70211 mus musculu
8	124	6.0	1070	1 PTK7_HUMAN	O13308 homo sapien
9	120	5.8	345	1 OPCM_HUMAN	Q14982 homo sapien
10	119	5.8	345	1 OPCM_BOVIN	P11834 bos taurus
11	118	5.7	345	1 OPCM_RAT	P32736 rattus norv
12	117.5	5.7	1443	1 NEOL_CHICK	Q90610 gallus gall
13	116	5.6	4393	1 PGBM_HUMAN	P98160 homo sapien
14	114.5	5.6	1461	1 NEOL_HUMAN	Q92859 homo sapien
15	113.5	5.5	416	1 RAGE_BOVIN	Q28173 bos taurus
16	113.5	5.5	1377	1 NEOL_RAT	P97603 rattus norv
17	112	5.4	764	1 ICCR_MOUSE	Q68180 mus musculu
18	110.5	5.4	524	1 BUTY_MOUSE	Q62556 mus musculu
19	110.5	5.4	620	1 SMP_COTIA	Q92154 coturnix co
20	109.5	5.3	249	1 CSF_DROME	P78310 homo sapien
21	107	5.2	365	1 CXAR_HUMAN	Q05793 mus musculu
22	107	5.2	3707	1 PGBM_MOUSE	P15364 dirosophila
23	106.5	5.2	333	1 ANML_DROME	Q99795 homo sapien
24	106	5.1	319	1 A33_HUMAN	P36333 xenopus lae
25	106	5.1	1092	1 NCAL_XENLA	P35329 mus musculu
26	104.5	5.1	862	1 FRRP_RAT	Q62786 rattus norv
27	104.5	5.1	879	1 LAMP_CHICK	Q98919 gallus gall
28	103.5	5.0	338	1 TYO3_HUMAN	Q06418 homo sapien
29	103.5	5.0	890	1 CD2_MOUSE	P08921 rattus norv
30	102.5	5.0	344	1 KIL0_RAT	Q92018 rattus norv
31	102.5	5.0	365	1 CXAR_MOUSE	P97792 mus musculu
32	102.5	5.0	2499	1 MPRL_BOVIN	P08169 bos taurus

34	102	5.0	880	1 TYO3_RAT	P55146 rattus norv
35	100.5	4.9	873	1 FAS2_DROME	P34082 dirosophila
36	100	4.9	1241	1 PBN_HUMAN	O60500 homo sapien
37	99.5	4.8	261	1 KLR8_RAT	P36374 rattus norv
38	99.5	4.8	1260	1 CAML_MOUSE	P11627 mus musculu
39	99.5	4.8	1906	1 KMLS_CHICK	P11799 gallus gall
40	99.5	4.8	2481	1 UN52_CAEEL	O06561 caenorhabdi
41	99	4.8	338	1 LAMP_HUMAN	Q13449 homo sapien
42	99	4.8	338	1 LAMP_RAT	Q62813 rattus norv
43	99	4.8	359	1 LACH_DROME	Q24372 dirosophila
44	99	4.8	830	1 SREC_HUMAN	O14162 homo sapien
45	99	4.8	880	1 TYO3_MOUSE	P55144 mus musculu

ALIGNMENTS

RESULT 1
RAGE_HUMAN STANDARD: PRT: 404 AA.
ID RAGE_HUMAN
AC Q15109; Q15279; Q9Y3R3; Q9H2X7;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Advanced glycosylation end product precursor
DE (Receptor for advanced glycosylation end products).
GN AGER OR RAGE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=92340547; PubMed=1378843;
RA Neeser M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
RA Elliston K., Stern D., Shaw A.;
RT "Cloning and expression of a cell surface receptor for advanced
RT glycosylation end products of proteins."
RT J. Biol. Chem. 267:14998-15004(1992).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE=95137587; PubMed=7835890;
RX Suyaya K., Ikegawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,
RA Inoko H., Ikegawa T.;
RT "Three genes in the human MHC class III region near the junction with
RT the class II: gene for receptor of advanced glycosylation end
RT products, PBX2 homeobox gene and a notch homolog, human counterpart
RT of mouse mammary tumor gene int-3.";
RL Genomics 23:408-419(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Spies T., Hood L.;
RL Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.
RX Abeidin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.;
RA Yamamoto H.;
RT "Molecular heterogeneity of the receptor for advanced glycation
RT endproducts.";
RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX Mallerbe P., Richards J., Galliard H., Thompson A., Diener C.,
RA Schuler A., Huber G.;
RT "CDNA cloning of a novel secreted isoform of the human receptor for
RT advanced glycation end products (RAGE) and characterization of cells
RT co-expressing cell-surface scavenger receptors and Swedish mutant
RT amyloid precursor protein.";
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE-Lung;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-12 FROM N.A.
 RA Hudson B.I., Futers T.S.;
 RT "Novel polymorphisms in the receptor for advanced glycation
 end-products (RAGE) gene."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
 PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 RATE IN DIABETES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
 CC Secreted (isoform 2).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2/RAGESEC;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M91211; AA03574.1; -
 DR EMBL: D28769; BA05958.1; -
 DR EMBL: U89336; AA47491.1; -
 DR EMBL: AB036432; BA089369.1; -
 DR EMBL: AJ133822; CAB43108.1; -
 DR EMBL: BC020669; AA020669.1; -
 DR EMBL: AF208289; AA035728.1; -
 DR GenBank: HGNC:320; AGER.
 DR MIM: 600214; -
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_2.
 DR SMART: SMO0410; Ig_Like; 1.
 DR SMART: SMO0408; IGC2; 1.
 DR PROSITE: PS00290; IGC2; 1.
 DR Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
 KW Alternative splicing; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 404
 FT
 FT DOMAIN 23 342
 FT TRANSMEM 343 363
 FT DOMAIN 364 404
 FT DOMAIN 31 106
 FT DOMAIN 137 215
 FT DOMAIN 252 308
 FT DISULFID 38 99
 FT DISULFID 144 208
 FT DISULFID 259 301
 FT CARBOHYD 25 25
 FT CARBOHYD 81 81
 FT DOMAIN 380 384
 FT VARSPIC 54 67
 FT VARSPIC 275 404
 FT
 FT VARIANT 100 100
 FT
 FT CONFLICT 1 1
 M -> G (IN REF. 1).
 /FTID=VAR_011338.

SQ SEQUENCE 404 AA: 42802 MW: 0D584C436C30CCE7 CRC64:
 Query Match 7.2%; Score 147.5; DB 1; Length 404;
 Best Local Similarity 22.6%; Pred. No. 0.00012;
 Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;
 QY 15 EVIEGPNATVLKSGARFNCVVSOG---WKIMMALSDMYVLSRPMPEITTDRETS- 70
 DB 125 EVIDASSELTA---GVNKKGTCTVSESYAGTLSHLLG-----KPLVPEKGVSV 173
 QY 71 ----QRYDGGNT--SEMIINVEPSDGNIR---CSLONSRLHGSAYLTVOVMGELE 120
 DB 174 KEQTRRHPEPTGFTLOSELN---VTPARGDPRPTFCSCFSGPLRHHALTAIRQRRW 230
 QY 121 IP---SVNLVAENRP-----CEVCLSHHTRLPDISWELGLVSHSV 162
 DB 231 EAPVLEEVGLV---EPREGANAPGQVTLTCEVPAQS---PQIHMMD----- 274
 QY 163 YVPEPSDLSAVSLALTPQSGNGLTCVATWKSLSKRSATVNLTVRCPODTGGINI 222
 DB 275 -GVPLPPEPSVLLPEIGPODQGYSCVATHSSHGPOESRAVSITIE-PGEIG----- 327
 QY 223 PGVLSLSLSIGSLPTMGKVGIGLAGTMLP-----PCTLIIRCCCRRCGCCNC 275
 DB 328 -----PTASVSGSGSLGTLALGILGLGTLALIGVILQRR----- 366
 QY 276 CACCCCKRRKRFRIQFOKKSEKERT--NKEETESNGNSGVNSDEKTTVPASLPKS 333
 DB 367 -----ORGEERKAPENQEEERAEIN----- 389
 QY 334 CESSDPEQNRSSCGPP 349
 DB 390 -QSEPEAGSGESTGPP 404
 RESULT 2
 G55A_CHICK STANDARD; PRT; 337 AA.
 AC Q98892;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurite inhibitor GP35-A precursor (OBCAM protein gamma isoform).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA Lodge A.P., Kim D.-S., Howard M.R., McNamee C.J., Smith N., Moss D.J.;
 RT "Cloning of CEPU-s, a secreted isoform of CEPU-1, and OBCAM cDNAs from
 chick: structural diversity of IGLON family proteins."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 64-337 FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE=97157768; PubMed=9004047;
 RA Wilson D.J.A., Kim D.-S., Clarke G.A., Marshall-Clarke S., Moss D.J.;
 RT "A family of glycoproteins (GP55), which inhibit neurite outgrowth,
 are members of the Ig superfamily and are related to OBCAM,
 J. Cell Sci. 109:3129-3138(1996).
 CC -1- FUNCTION: INHIBITS NEURITE OUTGROWTH.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: INCREASES DURING DEVELOPMENT FROM VERY LOW
 LEVELS AT EMBRYONIC DAY 10 AND IS MOST ABUNDANT AFTER HATCHING.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 SUBFAMILY.
 CC -1- STABILITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: Y08170; CAB41420.1; -
InterPro: IPR003006; I9_MHC.
InterPro: IPR003598; I9_C2.
InterPro: IPR003600; I9_Like.
Pfam: PF00047; I9; 3.
SMART: SM00410; I9_Like; 1.
SMART: SM00408; I9_C2; 2.
Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
Repeat: Signal.
Signal: 1
CHAIN 1 20
PROPEP 21 317
DOMAIN 318 337
DOMAIN 43 115
DOMAIN 143 201
DOMAIN 229 295
DOMAIN 50 108
DISULFID 150 194
DISULFID 236 288
CARBOHYD 133 133
CARBOHYD 277 277
CARBOHYD 285 285
CARBOHYD 298 298
SEQUENCE 337 AA; 36887 MW; BAE71755185651E CRC64;

Query Match
Best Local Similarity 25.8%; Pred. No. 0.002; DB 1; Length 337;
Matches 65; Conservative 33; Mismatches 96; Indels 58; Gaps 14;

22 NATVLKGSQARFNCYVSGQKLMALSDMYL-----SYRPEPIITDRFSGYD 74
37 NATVROGESATLCTCYDDRRVRVAW-LNRSTLYAGNDKWSIDNRVILSN---TKIQY 91
75 QGNGFTSEMIHNVPSDSNGIRCSLO-----NSRLHGSAYLVYQVMEFLFIPSVNLV 128
92 -----SIKIHNVYDDESPYCSQVOTDNHPTSRVH---LIYQVPOIVNISSDITV 140
129 AENPECEVTCV---PSHWTRLPDISWELGLVSHSYFVPPPSDLOASVILATLPOS 184
141 NEGSYVILMCLAFGRPE---PTVWTR---HLGKGGQGVSDDEYLE---ITGITRQ 188
185 NGTLCTVATWKSILAKRSATVNLTV-----IRCPQDTGGGINIPGVL---SSLPSLGS 235
189 SGEYECASV-NDVAVPDVRYKVTYVNPYISNAKNTGASVQKGLIQCASAVPAVERQ 247
236 LPTMGVGLGLA 247
248 ---WKEEDTRLA 256

RESULT 3
NTRI RAT STANDARD: PRT; 344 AA.

AC 062718;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Neurotrophin precursor (GP65).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.
RC STRAIN-Sprague-Dawley;
RX MEDLINE=95138094; PubMed=7891157;

RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,
RT Salzer J.L.;
RT "Cloning of neurotrophin defines a new subfamily of differentially
RT expressed neural cell adhesion molecules."
RL Neurosci. 15:2141-2156(1995).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN SEVERAL
CC SUBLETA, AND LOWER CORTICAL LAMINAE IN THE FOREBRAIN AND IN THE
CC PONTINE NUCLEUS, CEREBELLAR GRANULE CELLS, AND PURKINJE CELLS IN
CC THE HINDRAIN.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON

CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

EMBL: U16845; AA67445.1; -
InterPro: IPR003006; I9_MHC.
InterPro: IPR003598; I9_C2.
InterPro: IPR003600; I9_Like.
Pfam: PF00047; I9; 3.
SMART: SM00410; I9_Like; 1.
SMART: SM00408; I9_C2; 2.
Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
Repeat: Signal.

CHAIN 1 31
PROPEP 32 321
DOMAIN 322 344
DOMAIN 50 122
DOMAIN 150 208
DOMAIN 236 302
DOMAIN 57 115
DISULFID 157 201
DISULFID 243 295
CARBOHYD 44 44
CARBOHYD 70 70
CARBOHYD 152 152
CARBOHYD 216 216
CARBOHYD 284 284
CARBOHYD 292 292
CARBOHYD 305 305
CARBOHYD 321 321
CARBOHYD 321 321
SEQUENCE 344 AA; 37998 MW; CBB39BE53B3B224 CRC64;

Query Match
Best Local Similarity 24.2%; Pred. No. 0.0028; DB 1; Length 344;
Matches 64; Conservative 43; Mismatches 104; Indels 53; Gaps 14;

4 VIFLHSGS---SGNEVI-EGPQNAVYKGSQARFNCYVSGQKLMALSDMYLVS--- 55
21 LFLVPTGVPRVSGDAPPKAMDNTVTRQGSATLCTIINDRYRVRVAMLRSTLYAGND 80
56 ---VRMEPIITNDRTSQRVDOGNFTSEMIHNVPSDSNGIRCSLO-----NSRLH 106
81 KWLCDRLVYVLSN---TQIQY-----SIKIHNVYDDESPYCSQVOTDNHPTSRVH 129
107 GSAVYLVYQVMEFLFIPSVNLTVVAENPECEVTCVPSHWTRLPDISWELGLVSHSYFV 166
130 ---LIYQVSPKIVELISDISINEGNISLCTAIGRPE-PTVWTR---HISPRKAVGVS 181
167 EPPSDQSAVSLALTPQSNCTLCTVATWKSILAKRSATVNLTVICP-----QDTGGGIN 221
182 EDEYLE---IQGITRQSGEYECASV-NDVAVPVYRVRVNVNPYISNAKNTGASVQKGLI 236

FT	CARBOHD	282	282	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD	290	290	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD	303	303	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID	330	330	GPI-ANCHOR (POTENTIAL).
FT	VARSPLIC	310	320	MISSING (IN MAJOR ISOFORM).
SEQ	SEQUENCE	353 AA;	38736 MM;	2550C48591EBBBA6 CNC64;

Query Match	Best Local Similarity	6.28;	Score 128;	DB 1;	Length 353;
Matches 63;	Conservative 36;	Mismatches 90;	Indels 56;	Gaps 13	

ID	DCG_HUMAN	STANDARD;	PRT: 1447 AA.
AC	P43146;		
AD	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
GN	Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).		
GN	DCC.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homindaes; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95011532; PubMed=7926722;		
RA	Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,		
RA	Vogelstein B.;		
RT	"The DCC gene product in cellular differentiation and colorectal		
RT	tumorigenesis."		
RT	Genes Dev. 8:1174-1183(1994).		
RL	[2]		
RP	SEQUENCE OF 1-750 FROM N.A.		
RX	MEDLINE=90100559; PubMed=2294591;		
RA	Featon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,		
RA	Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,		
RA	Vogelstein B.;		
RT	"Identification of a chromosome 10q gene that is altered in		
RT	colorectal cancers."		
RT	Science 247:49-56(1990).		
RL	[3]		
RP	SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).		
RX	MEDLINE=91121517; PubMed=199132;		
RA	Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,		
RA	Oliver J.D., Kinzler K.W., Vogelstein B.;		
RT	"Scrambled exons."		
RT	Cell 64:607-613(1991).		
RL	[4]		
RP	GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.		
RX	MEDLINE=94245241; PubMed=8188295;		
RA	Cho K.R., Oliner J.D., Simons J.W.,		

Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.,
 "The DCC gene: structural analysis and mutations in colorectal
 carcinomas." *Cell* 58:287-292 (1989).
 RT Genomics 19:525-531(1994).
 [5]
 RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.
 RX MEDLINE=94243823; PubMed=8187090;
 RA Miyake S., Nagai K., Yoshino K., Ota M., Endo M., Yasa Y.,
 "Point mutations and allelic deletion of tumor suppressor gene DCC in
 human esophageal squamous cell carcinomas and their relation to
 metastasis." *Cancer Res.* 54:3007-3010(1994).
 RT Cancer Res. 54:3007-3010(1994).
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X76132; CAAS3735.1; -
 DR EMBL; M32292; AAA53751.1; -
 DR EMBL; M32286; AAA52174.1; -
 DR EMBL; M32288; AAA52175.1; ALT-SEQ.
 DR EMBL; M32290; AAA52176.1; -
 DR EMBL; M63696; AAA52177.1; -
 DR EMBL; M63700; AAA52178.1; -
 DR EMBL; M63702; AAA52179.1; -
 DR EMBL; M63718; AAA52180.1; -
 DR EMBL; M63698; AAA52181.1; -
 DR PIR; A54100; A54100.
 DR PIR; A40098; A40098.
 DR PIR; A38442; A38442.
 DR HSSP; PS6276; 1TLK.
 DR Genew; HGNC:2701; DCC.
 DR MIM; 120470; -
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003962; FNIII_repeat.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003600; IG_II-like.
 DR Pfam; PF00047; Ig_4.
 DR PRINTS; PR00014; FNTYPEIII.
 DR SMART; SM00060; FN3; 6.
 DR SMART; SM00410; IG_II-like; 2.
 DR SMART; SM00408; IGC2; 3.
 DR Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.
 KM SIGNAL 1 25
 FT CHAIN 26 1447
 FT DOMAIN 26 1097
 FT TRANSMEM 1098 1122
 FT DOMAIN 1123 1447
 FT DOMAIN 154 124
 FT DOMAIN 154 219
 FT DOMAIN 254 317
 FT DOMAIN 345 407
 FT DOMAIN 426 522
 FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.

Query Match
 Best Local Similarity 24.8%; Pred. No. 0.021; 99; Indels 50; Gaps 10;
 Matches 60; Conservative 33; Mismatches 99; Length 1447;
 Score 128; DB 1;
 Query 10 SGGSGNEV-----IEGPNATVLSGQAFNCTVSGW--KLIMWALSDM 51
 Db 220 SRTGNEARVRLSPGHLRDTFLQRPNSVVALEGDAVLECVCV-GYPPSPFTYLRGEE 278
 QY 52 VLVSRPPEPITITDRFTSQRDGGNFTSEMIITHNVEPDSGNIC--SLQSRFLHGA 109
 Db 279 VI-----QLSKKYSILGG--SNLISNTVDDSGWTCVYTKKENSASA 323
 QY 110 VLTVOYGEELFIPSVNLVVAENPECEVTCPSHTRLPDISM-ELGLVSHSSTYVPEP 168
 Db 324 ELTVLPPEPFLNHPNSNLVAYESMDIEFCYVS-GKPVPTVMMKMGDVPISDYFQIVGG 382
 QY 169 SDQSASVITALTPQSNGLTFCVATWKLARKSNVTLVYIRPODPTGGINIPVLSS 228
 Db 383 SNLR-----ILGVVKSDEGFCQVAENEGNAOTSQTLVPEKPAIPSSS-----VLPS 430
 QY 229 LP 230
 Db 431 AP 432
 RESULT 6
 KMLS_HUMAN STANDARD; PRT; 1914 AA.
 ID KMLS_HUMAN Q15746; Q14844; Q16794; Q95796; Q95797; Q95798; Q95799; Q9UBG5;
 AC Q9UBG5;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin light chain kinase, smooth muscle and non-muscle isozymes
 DE (EC 2.7.1.117) (MCK) [Contains: Yelokin (Kinase related protein)
 DE (KRP)]
 DE GN Homo sapiens (Human).
 OS Mylk OR MCK.
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=umbilical vein endothelial cells;
 RX MEDLINE=97304466; PubMed=9160829;
 RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,
 Verin A.D.;

RT "Myosin light chain kinase in endothelium: molecular cloning and
 regulation.";
 RL Am. J. Respir. Cell Mol. Biol. 16:489-494(1997).
 RN [2]
 RP REVISIONS.
 RA Birukov K.G., Garcia J.G.N.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).
 RC TISSUE-Umbilical vein;
 RA Lazar V.L., Garcia J.G.N.;
 RL "A single human myosin light chain kinase gene (MLCK; MYLK).";
 RN Genomics 57:256-267(1999).
 RP REVISIONS (ISOFORM 2).
 RA Birukov K.G., Garcia J.G.N.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 923-1914 FROM N.A.
 RC TISSUE-Hippocampus;
 RA MEDLINE-96121365; PubMed-855746;
 RL Potter M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossier J.,
 RA Turnell W.G.;
 RL "The human myosin light chain kinase (MLCK) from hippocampus:
 cloning, sequencing, expression, and localization to 3qcen-q21.";
 RN Genomics 29:562-570(1995).
 RP [6]
 RC SEQUENCE OF 1614-1914 FROM N.A.
 RA TISSUE-Lung, and Placenta;
 RL MEDLINE-2007838; PubMed-10536370;
 RA Waterston D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,
 RA Shrinley V.P., Van Eldik L.J., Halech J.;
 RL "Analysis of the kinase-related protein gene found at human chromosome
 3q21 in a multi-gene cluster: organization, expression, alternative
 J. Cell. Biochem. 75:481-491(1999).
 RN [7]
 RP SEQUENCE OF 1456-1914 FROM N.A.
 RC TISSUE-Placenta;
 RL Waterston M.D.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [1]
 RP FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR
 SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE
 IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT
 DETERMINES THE INTERACTION WITH ACTIN FILAMENTS. CENTRAL
 DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE
 REMODELING. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO
 CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE
 AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED
 BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT
 IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.
 CC CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
 light-chain] phosphate.
 CC SUBUNIT: TELOKIN BINDS CALMODULIN.
 CC ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: A NON-MUSCLE FORM (THE
 LONGEST FORM), A SMOOTH-MUSCLE FORM AND TELOKIN (A C-TERMINAL
 INITIATION).
 CC ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: 1 (SHOWN HERE), 2, 3A,
 3B, 4 AND DEL-1790; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES ARE
 EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN
 CULTURED ENDOTHELIAL WITH QUALITATIVE EXPRESSION APPEARING TO BE
 THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN
 IS THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES.
 CC PPM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.
 CC SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U48959; AAC18423.2; -
 DR EMBL: AF096601; AAD15921.2; -
 DR EMBL: AF096602; AAD15922.1; -
 DR EMBL: AF096603; AAD15923.1; -
 DR EMBL: AF096604; AAD15924.1; -
 DR EMBL: X85337; CAA59685.1; -
 DR EMBL: AF096771; AAD51380.1; -
 DR EMBL: AF096766; AAD51380.1; JOINED.
 DR EMBL: AF096767; AAD51380.1; JOINED.
 DR EMBL: AF096768; AAD51380.1; JOINED.
 DR EMBL: AF096769; AAD51380.1; JOINED.
 DR EMBL: AF096770; AAD51380.1; JOINED.
 DR EMBL: AF096771; AAD51381.1; -
 DR EMBL: AF096774; AAD51381.1; -
 DR EMBL: AF096775; AAD51381.1; -
 DR EMBL: AF096776; AAD51381.1; JOINED.
 DR EMBL: X90870; CAA62378.1; -
 DR HSP: O63450; 1A06.
 DR Genew: HGNC:7590; MYLK.
 DR MIM: 600922; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003600; IG_Like.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF00047; Ig_8.
 DR Pfam: PF00069; pkinase_1.
 DR PRINTS: PR00014; FMYPEPIT.
 DR Prodom: PD000001; Euk_pkinase_1.
 DR SMART: SM00060; FN3_1.
 DR SMART: SM00410; IG_Like_1.
 DR SMART: SM00408; IG_C2_8.
 DR PROSITE: PS00107; S_TKc_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM_1.
 DR Transfaser: Serine/threonine-protein kinase; Calmodulin-binding;
 KW ATP-binding; phosphorylation; Immunoglobulin domain; Repeat;
 KW Alternative initiation; Alternating splicing.
 FT CHAIN 1 1914
 FT CHAIN 923 1914
 FT CHAIN 1761 1914
 FT INIT_MET 923 923
 FT INIT_MET 1761 1761
 FT DOMAIN 1343 1413
 FT DOMAIN 1464 1719
 FT DOMAIN 1711 1774
 FT DOMAIN 1824 1891
 FT NP_BIND 1470 1476
 FT BINDING 1493 1493
 FT AC_1_SITE 1585 1585
 FT AC_1_SITE 1585 1585
 FT DOMAIN 1906 1914
 FT DOMAIN 868 998
 FT REPEAT 868 895
 FT REPEAT 895 923
 FT REPEAT 924 951
 FT REPEAT 952 976
 FT REPEAT 980 998
 FT DOMAIN 999 1063
 FT REPEAT 999 1003
 FT 1-1
 FT 1-2
 FT 1-3
 FT 1-4
 FT 1-5 (INCOMPLETE).
 FT 6 x 12 AA APPROXIMATE TANDEM REPEATS.
 FT 11-1 (INCOMPLETE).


```

FT REPEAT 1004 1015 II-2.
FT REPEAT 1016 1027 II-3.
FT REPEAT 1028 1039 II-4.
FT REPEAT 1040 1051 II-5.
FT REPEAT 1052 1063 II-6.
FT VARSPLIC 437 506 VSGIPKPEVAMFLEGPVRROQSEIEYEDAGSHYLCLLKA
FT RTRSGVSCSTASAGVSCSWTLQYVER -> G (IN
FT ISOFORM 2 AND ISOFORM 3B).
FT VARSPLIC 1433 1439 DEVEVSD -> MKRRCOT (IN ISOFORM 3A,
FT ISOFORM 3B AND ISOFORM 4).
FT VARSPLIC 1473 1546 GKFGQVRLVCKVDAFEKRAVWVLETVSGEL -> L
FT MISSING (IN ISOFORM 3A AND ISOFORM 3B).
FT MISSING (IN ISOFORM DEL-1790).
FT V -> M (IN REF. 5).
FT S -> P (IN REF. 3).
FT S -> A (IN REF. 5).
FT P -> A (IN REF. 5).
FT KPM -> BAH (IN REF. 5).
FT P -> L (IN REF. 3).
FT L -> P (IN REF. 5).
FT E -> D (IN REF. 3).
FT M -> I (IN REF. 3).
FT AADI5924).
FT A -> G (IN REF. 5).
FT L -> S (IN REF. 5).
FT T -> S (IN REF. 5).
FT V -> C (IN REF. 5).
FT S -> T (IN REF. 3).
FT I -> T (IN REF. 5).
FT A -> P (IN REF. 5).
FT G -> R (IN REF. 5).
FT GY -> D (IN REF. 3).
FT AADI5922/AADI5923/
FT AADI5924).

```

Query Match 6.2%; Score 128; DB 1; Length 1914;
 Best Local Similarity 20.3%; Pred. No. 0.03;
 Matches 86; Conservative 58; Mismatches 149; Indels 130; Gaps 19;

```

QY 20 PONTATLVKSGQARFNCVSGM--KLIMWALSDMWLVSVRMEIITNDPFGSRVDOGG 77
DB 39 PNLICKEGATKFEGRV--RGVPEPOVWVH-----RNGOPTISGGRFL--LDCGI 85
QY 78 NTSSEMIITHNVPSPDSGNIRCSLONSRLHGSAYLTVOV-----MGEL 119
DB 86 RGTFSVLVHAHVEDRGTCTCATNG--SGARQVTVELTVEGSAKOLGCPVVSKITLGD 143
QY 120 FI-----PSV-----NLVVAENE-----PCEVTCPLPSHWTIRLPDISWELG- 154
DB 144 FMSAVETRPISIMGECPPKATKLGVRVYVKEQOMRFSCKITGRPO-----POVTWLGKN 198
QY 155 LLYVSHSYFFVPPSDLOSASVIALTPQSNGLTFCVATWKSLSKARKSA-----TVN 206
DB 199 VPLQPARVSVSEKNGMO--VLEHGVNDQDVYTCIVVNSGKMSABELISGIDSAN 257
QY 207 LTVIRCPDGTGGI-----NIPGVLSLPLSGPLPTWKGKGLAGTMLTPTCTLTIR 261
DB 258 RSEVREKKNANDVRKEVTVISKESKLDL-----288
QY 262 CCCRRRCGCCNCCRCCKRRKGRFRIOFOKSEKTEKNETESGNN--SGYNS 318
DB 289 -----EAAASKNCSNCPQKSGPPMAANSOPPP--RESKLEKSDPRTAPQTP 336
QY 319 DEOKTDTASLPKSCSESSDPQRRNSCGPPHROADQ--PRPASHPOQSFNLASPE 374
DB 337 VLQTTSSITLQARVQ--PEPRAPGLAVLSPSGEKKRPPAPRPATEPTPQGLIGSD 393
QY 375 KVS 377
DB 394 VVS 396

```

RESULT 7

```

DCC_MOUSE ID DCC_MOUSE STANDARD; PRT; 1447 AA.
AC P70211;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor suppressor protein DCC precursor.
DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC STRAIN-BALB/c; TISSUE-Brain;
RC MEDLINE=96112625; PubMed=8570174;
RA Cooper H.M., Armes P., Brito J., Gad J., Wilks A.F.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RL gene (mdcc) and its expression in the developing mouse embryo.";
RN Oncogene 11:2243-2254(1995).
RP REVISIONS.
RC STRAIN-BALB/c; TISSUE-Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
CC EXPRESSED DURING MID GESTATION. HIGHEST LEVELS
CC AND REMAIN AT THIS LEVEL IN THE ADULT.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: X85788; CAA59786.1; -.
DB HSSP: P56276; 1PLK.
DB MGD: MGI:94869; Dcc.
DB InterPro: IPR003961; FN.III.
DB InterPro: IPR003962; FN.III.repeat.
DB InterPro: IPR003006; I9_MHC.
DB InterPro: IPR003598; I9_C2.
DB InterPro: IPR003600; I9_Like.
DB Pfam: PF00041; fn3; 6.
DB Pfam: PF00047; I9; 4.
DB PRINTS: PRO0014; FNTPETII.
DB SMART: SM00060; FN3; 6.
DB SMART: SM00410; IG_Like; 2.
DB SMART: SM00408; IG_C2; 3.
DB GlycoProtein: Immunoglobulin domain; Transmembrane; Signal;
KW Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1447 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM A.
FT CHAIN 85 1447 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM B.
FT INT_MET 85 1447 FOR ISOFORM B.
FT DOMAIN 26 1097 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1098 1122 POTENTIAL.
FT DOMAIN 1123 1447 CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 54 124 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 154 219 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 254 317 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 345 407 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.
 FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.
 FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.
 FT DISULFID 61 117 BY SIMILARITY.
 FT DISULFID 161 212 BY SIMILARITY.
 FT DISULFID 261 310 BY SIMILARITY.
 FT DISULFID 352 400 BY SIMILARITY.
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLOT 819 838 MISSING (IN ISOFORM C).
 SQ SEQUENCE 1447 AA; 158298 MW; ODIF1097C22D5B9F CRC64;

Query Match 6.2%; Score 127; DB 1; Length 1447;
 Best Local Similarity 26.6%; Pred. No. 0.026; Matches 51; Conservative 29; Mismatches 86; Indels 26; Gaps 8;

DB 17 IEGPQNAVYKSGARFPCYSGW--KLIMWALSDWVLSVRMEPIITNDRFTSORVD 74
 DB 243 IORPSNVATIEGKQVNLDCVVS-GYPPSPFWLGEIEI-----QLRSKYS 288
 QY 75 OGNETSEMIIHNEPDSGNIRC--SLONSRLHGSXLYVOVGELFISVNLVAENE 132
 DB 289 ILGG--SLLLSNTTDDSGTYCTVYTKENISASALVLPVPLNHPMLVAYEES 346
 QY 133 PCEVTCUPSHWTRLPDJSW-ELGLVSHSSYTFPEPSDQSAVSIALATLPQSGNLTGV 191
 DB 347 DIEECAYS-GKPPVTVMMKMGNDVVPDYFOVGSNLR-----ILGVKSGEGYQCV 401
 QY 192 ATWKSILARKSA 203
 DB 402 AENEGAGNOSA 413

RESULT 8

PK7_HUMAN STANDARD; PRT; 1070 AA.

AC Q13308; Q13417;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase-like 7 precursor (Colon carcinoma kinase-4)
 GN (CCK-4).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon carcinoma, and Placenta;
 RX MEDLINE=96074849; PubMed=7478540;
 RA Mossie K., Jallat B., Alves F., Sures I., Plozman G.D., Ullrich A.;
 RT "Colon carcinoma kinase-4 defines a new subclass of the receptor
 tyrosine kinase family";
 RL Oncogene 11:2179-2184(1995).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=97037064; PubMed=8882711;
 RA Park S.-K., Lee H.-S., Lee S.-T.;
 RT "Characterization of the human full-length PK7 cDNA encoding a

FT receptor protein tyrosine kinase-like molecule closely related to
 FT chick KLG.";
 RT J. Biochem. 119:235-239(1996).
 RN [3]
 RP SEQUENCE FROM N.A., AND REVISION TO 834.
 RA Lee S.-T., Park S.-K., Lee H.-S., Ji A.R., Jung J.W.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY
 CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE
 CC PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR
 CC PROGRESSION MARKER.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, PANCREAS,
 CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID
 CC ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.
 CC - SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC - SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U3635; AAA87565.1; -
 CC EMBL; U40271; AAC50484.2; -
 CC EMBL; AF447157; AAL39062.1; -
 CC EMBL; AF447157; AAL39062.1; JOINED.
 CC EMBL; AF447158; AAL39062.1; JOINED.
 CC EMBL; AF447162; AAL39062.1; JOINED.
 CC EMBL; AF447164; AAL39062.1; JOINED.
 CC EMBL; AF447167; AAL39062.1; JOINED.
 CC EMBL; AF447170; AAL39062.1; JOINED.
 CC EMBL; AF447171; AAL39062.1; JOINED.
 CC EMBL; AF447173; AAL39062.1; JOINED.
 CC EMBL; AF447174; AAL39062.1; JOINED.
 CC EMBL; AF447175; AAL39062.1; JOINED.
 CC HSSP; P08631; IADS.
 CC Genew; HGNC:9618; PK7.
 DR MTM; 601890; -
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR002011; RtkinaseII.
 DR InterPro; IPR001245; Tyr_Pkinase.
 DR Pfam; PF00047; Ig; 7.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000001; Euk_Pkinase; 1.
 DR SMART; SM00410; IG_Like; 2.
 DR SMART; SM00219; TyrcK; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; FALSE NEG.
 KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 1070
 FT CHAIN 31 704
 FT DOMAIN 31 704
 FT TRANSMEM 726 1070
 FT DOMAIN 143 208
 FT DOMAIN 239 308
 FT DOMAIN 336 398
 FT DOMAIN 426 488
 FT DOMAIN 517 577
 FT POTENTIAL.
 FT TYROSINE-PROTEIN KINASE-LIKE 7.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE DOMAIN 1.
 FT IG-LIKE C2-TYPE DOMAIN 2.
 FT IG-LIKE C2-TYPE DOMAIN 3.
 FT IG-LIKE C2-TYPE DOMAIN 4.
 FT IG-LIKE C2-TYPE DOMAIN 5.
 FT IG-LIKE C2-TYPE DOMAIN 6.

FT DOMAIN 606 671 IG-LIKE C2-TYPE DOMAIN 7.
 FT PROTEIN KINASE; INACTIVE.
 FT DISULFID 796 1066
 FT DISULFID 53 101
 FT DISULFID 150 200
 FT DISULFID 246 301
 FT DISULFID 343 391
 FT DISULFID 433 481
 FT DISULFID 524 570
 FT DISULFID 613 664
 FT DISULFID 116 115
 FT CARBOHYD 175 175
 FT CARBOHYD 184 184
 FT CARBOHYD 214 214
 FT CARBOHYD 268 268
 FT CARBOHYD 283 283
 FT CARBOHYD 405 405
 FT CARBOHYD 463 463
 FT CARBOHYD 567 567
 FT CARBOHYD 646 646
 FT CONFLICT 92 92
 FT CONFLICT 147 147
 FT CONFLICT 207 207
 FT CONFLICT 495 496
 FT CONFLICT 515 515
 FT CONFLICT 881 881
 FT CONFLICT 969 969
 FT CONFLICT 992 992
 SQ SEQUENCE 1070 AA; 118260 MW; 47CDE25B8E3698A5 CRC64;

Query Match
 Best Local Similarity 25.3%; Score 124; DB 1; Length 1070;
 Matches 58; Conservative 32; Mismatches 103; Indels 36; Gaps 10;

QY 16 VIEGQNFVYKSGARPCNTVS-QGKLMALSDMYLSVREMEPTITNDRFTSQRD 74
 DB 227 VVLAPODVVAYREEMHCOFSPQSLQWLEFEDPTIRNRPRLRAVFA---- 282
 QY 75 OGNFTSEMIHNEPSSGNIRCSLQNSR-----LHSAVLTVQVNGELFIPSVNLVVA 129
 DB 283 -----NGSILLTVQVPRNAGIRYICIGOGRPITLLEATLHAELEDMPLFEPRVTFAGS 337
 QY 130 ENRPECVYTCLEPSHWRRLPDISWE-LGL-LVSHSYFYFPEPSDQSAVSIALTPQSN-G 186
 DB 338 EE---RVTCLEPKGLPEPSVMEHAGVRLPTGRVY-----QKGHELVLANIAESDAG 387
 QY 187 TLTCVATMKSUKARKSATVNLTV-----IRPOPTGGGINIPGLSSL 229
 DB 388 VYTCMAA-NLAGORQDVNITVATVPWMLKKRQDSQLEEGKPGILDCL 434

RESULT 9
 OPMC_HUMAN STANDARD; PRT; 345 AA.
 ID OPMC_HUMAN
 AC 014982;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OBPCAM)
 DE (Opioid-binding cell adhesion molecule) (OPCML).
 GN OPCML (Human)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Occipital cortex;
 RX MEDLINE=95237612; PubMed=7721093;
 RA Shark K.B., Lee N.M.;
 RT "Cloning, sequencing and localization to chromosome 11 of a cDNA
 encoding a human opioid-binding cell adhesion molecule (OBPCAM)."
 RL Gene 155:213-217(1995).
 CC -!- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY

CC INVOLVED IN CELL CONTACT
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L34774; AAA6387.1; -
 CC Genew; HGNC:8143; OPCML.
 CC MIM; 600632; -
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003598; Ig_C2.
 CC InterPro; IPR003600; Ig_Like.
 CC Pfam; PF00047; Ig_3.
 CC SMART; SM00408; IgC2; 2
 CC SMART; SM00410; Ig_Like; 1.
 CC DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;

DR Repeat; Signal.
 DR SIGNAL 1 27
 FT CHAIN 28 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULFID 57 115
 FT DISULFID 157 202
 FT DISULFID 244 296
 FT CARBOHYD 44 44
 FT CARBOHYD 70 70
 FT CARBOHYD 140 140
 FT CARBOHYD 285 285
 FT CARBOHYD 293 293
 FT CARBOHYD 306 306
 FT LIPID 322 322
 SQ SEQUENCE 345 AA; 38007 MW; E7AD17BE1AA3FE4 CRC64;

Query Match
 Best Local Similarity 25.0%; Score 120; DB 1; Length 345;
 Matches 71; Conservative 37; Mismatches 102; Indels 74; Gaps 18;

QY 4 VIFLHSG-----SGNEVT-EGPQNFVYKSGARPCNTVSQGKLMALSDMYVL----- 54
 DB 21 LLEFLVPTGVPRSGDATFPFANDNTVROGESATLRCTIDDEVRVAVM-LNRSTILVAGN 79
 QY 55 ---SVRMEPIITNDRFTSQRYDQGNFTSEMIHNEPSSGNIRCSLQ-----NSRL 105
 DB 80 DKMSIDPRVILIVN---TPYQY-----SIMQVVDVYDEGPYCSVQDTHNEKTSRV 128
 QY 106 HGSAYLVTVQVNGELFIPSVNLVVAENRPECVYTC-----PSHMTRLPDLSWELGLVSHS 161
 DB 129 H-----LTVQVPPQIMNISSDITVNEGSSVTLCLLAIGREP-----PTVTWR-----HLS 173
 QY 162 YV-----FVPEPSDQSAVSIALTPQSNGLTVCVATMKSUKARKSATVNLTV-----IRC 212
 DB 174 VKEGGQFVSEBYLE-----ISDIKRDQSGYECNAL-NDVAPADVRKVVITVAVPYISK 228
 QY 213 PDPTGGGINIPGVL-----SLPSLGFSLPWRGK-----VGLAG 248
 DB 229 AKNTGVSVGOKGILSCENASAVPAKFEQ---WFKETRLAIGDLOG 269

RESULT 10
 OPMC_BOVIN STANDARD; PRT; 345 AA.
 ID OPMC_BOVIN

AC P11834;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OB CAM)
 DE (Opioid-binding protein/cell adhesion molecule precursor (OB CAM))
 GN OPMCL OR OB CAM OR OCAM.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=89251576; PubMed=2721489;
 RA Schotfield P.R., McFarland K.C., Hayflick J.S., Wilcox J.N.,
 RA Cho T.M., Roy S., Lee N.M., Loh H.H., Seeburg P.H.;
 RT Molecular characterization of a new immunoglobulin superfamily
 RT protein with potential roles in opioid binding and cell contact.;
 RL EMBO J. 8:489-495(1989).
 CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY
 CC INVOLVED IN CELL CONTACT.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12672; CAA31192.1; -
 DR PIR: S03199; S03199; -
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR003600; Iq_Like.
 DR Pfam: PF00047; Iq; 3.
 DR SMART: SM00410; Iq_Like; 1.
 DR SMART: SM00408; IqC2; 2.
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 KW Repeat; Signal;
 FT SIGNAL 1 27
 FT CHAIN 28 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULFID 57 115
 FT DISULFID 157 202
 FT DISULFID 244 296
 FT CARBOHYD 44 44
 FT CARBOHYD 70 70
 FT CARBOHYD 140 140
 FT CARBOHYD 285 285
 FT CARBOHYD 293 293
 FT CARBOHYD 306 306
 FT LIPID 322 322
 SO SEQUENCE 345 AA; 37914 MW; D1ECC8D9E7D8CB19 CRC64;
 Query Match 5.88; Score 119; DB 1; Length 345;
 Best Local Similarity 25.18; Pred. No. 0.02; Mismatches 103; Indels 70; Gaps 17;
 Matches 70; Conservative 36; Mismatches 103; Indels 70; Gaps 17;
 QY 4 VIFHGS- - - - -SGNEVY-EGPONATVYKGSQARFCTVSGQKRLIMALSDMVYV- - - - -54
 Db 21 LFLVPTGVPRVSGDATPRKAMNVTVROGESATLRCCTIDRTYTRAV- - - - -LNSTILYKGN 79

QY 55 - - -SVREPEITITNDRTSORYDOGNFTSEMIHHVPSDSNINCSIQ- - - - -NSRL 105
 Db 80 DKMSIDPRVITLVN- - -TPQY- - - - -SINQVNDVYDESPYCSVQTNHPRKTSY 128
 QY 106 HGSAYLVTVQMGELFIPSVNLVAENEPCEVYCL- - -PSHWTRLPDISEWELGLVSHS 161
 Db 129 H- - - - -LIVQVPPQIMNISSDVYTNESSVYTLCLAGRE- - - - -PYTW- - - - -HLS 173
 QY 162 YV- - - - -FVPEPSDLOSASVSIATLTPQNSGTLCVATMSLRKRSATVNLVY- - - - -TRC 212
 Db 174 VKESQGVSEDEYLE- - - - -ISDIKROSEYECSL-NDVAPADVRKVKITVNPYYSK 228
 QY 213 PDGTGGINIPGL- - - - -SSLPSLGSLEPTGKVGIGLA 247
 Db 229 AKRTVSGVQKGLSCENASAVPMARFQ- - - - -WEKEDTRLA 264
 RESULT 11
 OPMCL_RAT
 ID OPMCL_RAT STANDARD: PRT; 345 AA.
 AC P32736; 001654; P32735; 001653;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OB CAM)
 DE (Opioid-binding protein/cell adhesion molecule precursor (OB CAM))
 GN OPMCL OR OB CAM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92347701; PubMed=1339369;
 RA Lipman D.A., Lee N.M., Loh H.H.;
 RT "Opioid-binding cell adhesion molecule (OB CAM)-related clones from a
 RT rat brain cDNA library.";
 RL Gene 117:249-254(1992).
 RN [2]
 RP SEQUENCE OF 195-214, AND GPI-ANCHOR.
 RX MEDLINE=95198094; PubMed=7891157;
 RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,
 RA Salzer J.L.;
 RT "Cloning of neurotrophin defines a new subfamily of differentially
 RT expressed neural cell adhesion molecules.";
 RL J. Neurosci. 15:2141-2156(1995).
 CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY
 CC INVOLVED IN CELL CONTACT.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;
 CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M88710; AAA40859.1; -
 DR EMBL: M88711; AAA40860.1; -
 DR EMBL: M88709; AAA40858.1; -
 DR PIR: JCI239; JCI239.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR003600; Iq_Like.
 DR Pfam: PF00047; Iq; 3.

DR SMART: SM00410; IG-like: 1.
 DR SMART: SM00408; IGC2: 2.
 KW Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
 KW Repeat: Signal; Alternative splicing.
 FT SIGNAL 1 27
 FT CHAIN 28 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULFID 57 115
 FT DISULFID 157 202
 FT CARBOHYD 244 296
 FT CARBOHYD 70 140
 FT CARBOHYD 140 285
 FT CARBOHYD 285 293
 FT CARBOHYD 306 306
 FT CARBOHYD 322 322
 FT LIPID 1 27
 FT VARSPLIC 1 27
 FT SEQUENCE 345 AA; 38067 MW; A3181B0753F9658E CRC64;
 Query Match 5.7%; Score 118; DB 1; Length 345;
 Best Local Similarity 24.7%; Pred. No. 0.024;
 Matches 71; Conservative 38; Mismatches 104; Indels 74; Gaps 18;
 4 VIFLGSG---SGNEVI-EGPQNAVYKSGQAFNCTVSQGRKLIMMALSDMVV----54
 21 LFLVPTGVPVRSGDAPFKAMDVTVQSGSATLRCITDRYTRAV-LNRSTIILACN 79
 55 ---SVRPEPITNRFORSRDGNGNTSEMIHNEVSDSGNIRSLQ-----NSRL 105
 80 DKMSIDPRVILVN---TPPTQ-----SIMIONDVDEGPTGCVQDNHPTSRV 128
 106 HGSAYLVVQVMEGLFIPSVNLVAANEPCVETCL---PSWTRLPDISMELGLVSHSS 161
 129 H-----LIVQVPPQIMNISDITVNEISSVTLICLAGRPE-----PLYTMR-----HLS 173
 162 YV---FVPEPSDLSAVSIILATLPQSGNGLTCVATWKSLSKRSATVNLTV-----TRC 212
 174 VKEGQFVSEDEYLE---ISDIKRDQSEFYECSSAL-NDVAPADVRAKVIHNPYPIYSK 228
 213 POGTGGGINIPGVL---SLPSLIGFSLPTWCK---VGLGLAGTML 251
 229 AKNTGVSVGOKGILSCASAVPMAFRO---WEKEDTRLATGLDGVRI 272
 RESULT 12
 NEOL-CHICK STANDARD; PRT: 1443 AA.
 ID NEOL-CHICK
 AC 090610:
 AC 16-OCT-2001 (Rel. 40, Created)
 AC 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neogenin (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus
 OK NCBI_taxid=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=embryonic brain;
 RX MEDLINE=95105243; PubMed=7806578;
 RA Vlemler J., Roman J.M., Dreyer W.J.;
 RT "Neogenin, an avian cell surface protein expressed during terminal
 neuronal differentiation, is closely related to the human tumor
 suppressor molecule deleted in colorectal cancer.";
 RT J. Cell Biol. 127:2009-2020(1994).
 RL
 CC -!- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE

CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
 CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
 CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -!- DEVELOPMENTAL STAGE: IN RETINA, EXPRESSED ON GANGLION CELL FIBERS.
 CC AS SOON AS THEY BEGIN TO EXTEND THEIR AXONS.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 007644; AAC59662.1; -
 CC HSSP: P11276; 2MFN.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003963; Ig_MHC.
 DR InterPro: IPR003964; Ig_C2.
 DR InterPro: IPR003965; Ig_Like.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; Ig; 4.
 DR PRINTS: PR00014; FNTYPEIII.
 DR SMART: SM00408; IGC2, 2.
 DR SMART: SM00410; IG-like: 2.
 DR SMART: SM00410; IG-like: 2.
 DR Transmembrane: 1
 DR NON_TER 1
 FT DOMAIN 1091 1111
 FT TRANSMEM 1091 1111
 FT DOMAIN 1112 1443
 FT DOMAIN 132 194
 FT DOMAIN 229 293
 FT DOMAIN 321 383
 FT DOMAIN 422 519
 FT DOMAIN 522 615
 FT DOMAIN 616 714
 FT DOMAIN 720 814
 FT DOMAIN 835 935
 FT DOMAIN 936 1037
 FT DISULFID 40 95
 FT DISULFID 139 187
 FT DISULFID 236 286
 FT DISULFID 328 376
 FT CARBOHYD 339 393
 FT CARBOHYD 176 176
 FT CARBOHYD 292 292
 FT CARBOHYD 456 456
 FT CARBOHYD 475 475
 FT CARBOHYD 625 625
 FT CARBOHYD 700 700
 FT CARBOHYD 894 894
 FT SEQUENCE 1443 AA; 158050 MW; 558C6795579C0E26 CRC64;
 Query Match 5.7%; Score 117.5; DB 1; Length 1443;
 Best Local Similarity 26.2%; Pred. No. 0.15;
 Matches 49; Conservative 26; Mismatches 73; Indels 37; Gaps 11;
 17 IEGPQNAVYKSGQAFNCTVSQGRKLIMMALSDMVVSVRPEPITNDRFTISQRY- 73
 218 VQPSLSLTKVQGNVAFPC-VAGGPPPYVKTNG-----EELITD---SERFA 264
 74 -DQGNFTSEMIHNEVSDSGNIRCL-SLNSRLHGSAYLVVQVMEGLFIPSVNLVAV 130
 265 LRAGGS---LITSDVTEDEVGTTCIADNENETEAQAEIAYGVVPEFLKRPANIYAH 320

OY 131 NE-----PCEVTCIPSHMTLPDISW-ELGLVSHSYFVEPESDLOSANSIALPEQSN 185
 Db 321 SMDIVECEVETGKPT-----PTVMKVNKGNDVYIPSYDFKVKENLQ-----VLGVKSPDE 371
 OY 186 GTLTGCA 192
 | | |
 Db 372 GYQOCIA 378

RESULT 13

PGBM_HUMAN

ID PGBM_HUMAN STANDARD; PRT; 4393 AA.

AC P98160; Q16287;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Basement membrane-specific heparan sulfate proteoglycan core

GN HSPG2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=92112994; PubMed=1730768;

RA Kallunki P., Tryggvason K.;

RA "Human basement membrane heparan sulfate proteoglycan core protein: a

RA 467-kD protein containing multiple domains resembling elements of the

RA low density lipoprotein receptor, laminin, neural cell adhesion

RA molecules, and epidermal growth factor.";

RL J. Cell Biol. 116:559-571(1992).

RN [2]

RP SEQUENCE FROM N.A.

RA TISSUE=Skin, and Colon;

RA MEDLINE=92235084; PubMed=1569102;

RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;

RA "Primary structure of the human heparan sulfate proteoglycan from

RA domains homologous to the low density lipoprotein receptor, laminin,

RA neural cell adhesion molecules, and epidermal growth factor.";

RL J. Biol. Chem. 267:8544-8557(1992).

RN [3]

RP SEQUENCE OF 1018-1472 FROM N.A.

RA TISSUE=Colon;

RA MEDLINE=91365376; PubMed=1679749;

RA Dodge G.R., Kovacs I., Chu M.L., Hassell J.R., McBride O.W.;

RA "Heparan sulfate proteoglycan of human colon: partial molecular

RA cloning, cellular expression, and mapping of the gene (HSPG2) to the

RA short arm of human chromosome 1.";

RL Genomics 10:673-680(1991).

RN [4]

RP SEQUENCE OF 892-1398 FROM N.A.

RA TISSUE=Fibrosarcoma;

RA MEDLINE=92120660; PubMed=1685141;

RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.;

RA "Cloning of human heparan sulfate proteoglycan core protein,

RA assignment of the gene (HSPG2) to 1p36.1--p33 and identification of

RA a BamHI restriction fragment length polymorphism.";

RL Genomics 11:389-396(1991).

RN [5]

RP SEQUENCE OF 1-21 FROM N.A.

RA MEDLINE=94052171; PubMed=8234307;

RA Cohen I.R., Graessels S., Murdoch A.D., Iozzo R.V.;

RA "Structural characterization of the complete human perlecan gene and

RA its promoter.";

RL Proc Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).

CC -1- FUNCTION: This protein is an integral component of basement

CC membranes. It is responsible for the fixed negative electrostatic

CC charge and is involved in the charge-selective ultrafiltration

CC properties. It serves as an attachment substrate for cells.

CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
 CC AND O-LINKED OLIGOSACCHARIDES.
 CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
 CC -1- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 22 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X62515; CAA44373.1; -

CC EMBL: M85289; AA52700.1; -

CC EMBL: M64283; AA52699.1; -

CC EMBL: S76436; AA52121.2; -

CC EMBL: L22078; - NOT_ANNOTATED_CDS.

CC HSPG2; P00740; IEDM.

CC Slenia-2DPAGE; P98160; -

CC GeneW: HGNC:5273; HSPG2.

CC MIM: 142461; -

CC InterPro: IPR000561; EGF-like.

CC InterPro: IPR000742; EGF-2.

CC InterPro: IPR001438; EGF-11.

CC InterPro: IPR003006; Iq_MHC.

CC InterPro: IPR003598; Iq_C2.

CC InterPro: IPR002172; LDL_recept_LA.

CC InterPro: IPR000034; Laminin_B.

CC InterPro: IPR002049; Laminin_EGF.

CC InterPro: IPR001791; Laminin_G.

CC InterPro: IPR000082; SEA_domain.

CC Pfam: PF00008; EGF_4.

CC Pfam: PF00047; Iq_22.

CC Pfam: PF00052; Laminin_B_3.

CC Pfam: PF00053; Laminin_EGF_7.

CC Pfam: PF00054; Laminin_G_3.

CC Pfam: PF00057; Iq_recept_La_4.

CC PRINTS: PR00190; SEA; 1.

CC PRODOM: PD003031; Laminin_B_3.

CC SMART: SM00180; EGF_Lam_6.

CC SMART: SM00001; EGF_Like_8.

CC SMART: SM00408; Iq_C2_22.

CC SMART: SM00192; LDLa_4.

CC SMART: SM00281; Lam_3.

CC SMART: SM00282; Lam_3.

CC PROSITE: PS00200; SEA; 1.

CC PROSITE: PS00022; EGF_1; 9.

CC PROSITE: PS01186; LAMININ_2; 5.

CC PROSITE: PS01248; LAMININ_TYPE_EGF_11.

CC PROSITE: PS50025; LAM_G_DOMAIN_3.

CC PROSITE: PS01209; LDLRA_1; 4.

CC PROSITE: PS50068; LDLRA_2; 4.

CC PROSITE: PS50024; SEA; 1.

CC Signal: Basement membrane: Proteoglycan; Repeat: Glycoprotein;

CC Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;

CC Extracellular matrix; EGF-like domain.

CC SIGNAL: 1 21 POTENTIAL.

CC CHAIN 22 4393 BASEMENT MEMBRANE-SPECIFIC HEPARAN

CC DOMAIN 80 194 SULFATE PROTEOGLYCAN CORE PROTEIN.

```

FT DOMAIN 197 236 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 283 321 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 323 361 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 366 405 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 405 506 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 523 532 LAMININ EGF-LIKE 1 (N-TERMINAL).
FT DOMAIN 533 732 LAMININ DOMAIN IV.1 (DOMAIN III A).
FT DOMAIN 733 765 LAMININ EGF-LIKE 1 (C-TERMINAL).
FT DOMAIN 766 815 LAMININ EGF-LIKE 2.
FT DOMAIN 816 823 LAMININ EGF-LIKE 3.
FT DOMAIN 825 925 LAMININ EGF-LIKE 4 (INCOMPLETE).
FT DOMAIN 881 935 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 926 1127 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 936 1127 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 1128 1160 LAMININ EGF-LIKE 6.
FT DOMAIN 1161 1210 LAMININ EGF-LIKE 7.
FT DOMAIN 1211 1267 LAMININ EGF-LIKE 8.
FT DOMAIN 1277 1326 LAMININ EGF-LIKE 9 (N-TERMINAL).
FT DOMAIN 1327 1336 LAMININ EGF-LIKE 9 (C-TERMINAL).
FT DOMAIN 1337 1531 LAMININ EGF-LIKE 9 (C-TERMINAL).
FT DOMAIN 1532 1564 LAMININ EGF-LIKE 10.
FT DOMAIN 1565 1614 LAMININ EGF-LIKE 11.
FT DOMAIN 1615 1672 LAMININ EGF-LIKE 11.
FT DOMAIN 1673 1773 LAMININ EGF-LIKE 11.
FT DOMAIN 1774 1867 LAMININ EGF-LIKE 11.
FT DOMAIN 1868 1957 LAMININ EGF-LIKE 11.
FT DOMAIN 1958 2053 LAMININ EGF-LIKE 11.
FT DOMAIN 2054 2153 LAMININ EGF-LIKE 11.
FT DOMAIN 2154 2246 LAMININ EGF-LIKE 11.
FT DOMAIN 2247 2342 LAMININ EGF-LIKE 11.
FT DOMAIN 2343 2438 LAMININ EGF-LIKE 11.
FT DOMAIN 2439 2535 LAMININ EGF-LIKE 11.
FT DOMAIN 2536 2631 LAMININ EGF-LIKE 11.
FT DOMAIN 2632 2728 LAMININ EGF-LIKE 11.
FT DOMAIN 2729 2828 LAMININ EGF-LIKE 11.
FT DOMAIN 2829 2926 LAMININ EGF-LIKE 11.
FT DOMAIN 2927 3023 LAMININ EGF-LIKE 11.
FT DOMAIN 3024 3114 LAMININ EGF-LIKE 11.
FT DOMAIN 3115 3213 LAMININ EGF-LIKE 11.
FT DOMAIN 3214 3300 LAMININ EGF-LIKE 11.
FT DOMAIN 3301 3401 LAMININ EGF-LIKE 11.
FT DOMAIN 3402 3490 LAMININ EGF-LIKE 11.
FT DOMAIN 3491 3576 LAMININ EGF-LIKE 11.
FT DOMAIN 3577 3664 LAMININ EGF-LIKE 11.
FT DOMAIN 3665 3845 LAMININ EGF-LIKE 11.
FT DOMAIN 3846 3883 LAMININ EGF-LIKE 11.
FT DOMAIN 3884 3924 LAMININ EGF-LIKE 11.
FT DOMAIN 3925 4105 LAMININ EGF-LIKE 11.
FT DOMAIN 4106 4143 LAMININ EGF-LIKE 11.
FT DOMAIN 4144 4178 LAMININ EGF-LIKE 11.
FT DOMAIN 4179 4203 LAMININ EGF-LIKE 11.
FT DOMAIN 4204 4391 LAMININ EGF-LIKE 11.
FT DOMAIN 4392 4478 LAMININ EGF-LIKE 11.
FT DOMAIN 4479 4515 LAMININ EGF-LIKE 11.
FT DOMAIN 4516 4553 LAMININ EGF-LIKE 11.
FT DOMAIN 4554 4591 LAMININ EGF-LIKE 11.
FT DOMAIN 4592 4629 LAMININ EGF-LIKE 11.
FT DOMAIN 4630 4667 LAMININ EGF-LIKE 11.
FT DOMAIN 4668 4705 LAMININ EGF-LIKE 11.
FT DOMAIN 4706 4743 LAMININ EGF-LIKE 11.
FT DOMAIN 4744 4781 LAMININ EGF-LIKE 11.
FT DOMAIN 4782 4819 LAMININ EGF-LIKE 11.
FT DOMAIN 4820 4857 LAMININ EGF-LIKE 11.
FT DOMAIN 4858 4895 LAMININ EGF-LIKE 11.
FT DOMAIN 4896 4933 LAMININ EGF-LIKE 11.
FT DOMAIN 4934 4971 LAMININ EGF-LIKE 11.
FT DOMAIN 4972 5010 LAMININ EGF-LIKE 11.
FT DOMAIN 5011 5048 LAMININ EGF-LIKE 11.
FT DOMAIN 5049 5086 LAMININ EGF-LIKE 11.
FT DOMAIN 5087 5124 LAMININ EGF-LIKE 11.
FT DOMAIN 5125 5162 LAMININ EGF-LIKE 11.
FT DOMAIN 5163 5200 LAMININ EGF-LIKE 11.
FT DOMAIN 5201 5238 LAMININ EGF-LIKE 11.
FT DOMAIN 5239 5276 LAMININ EGF-LIKE 11.
FT DOMAIN 5277 5314 LAMININ EGF-LIKE 11.
FT DOMAIN 5315 5352 LAMININ EGF-LIKE 11.
FT DOMAIN 5353 5390 LAMININ EGF-LIKE 11.
FT DOMAIN 5391 5428 LAMININ EGF-LIKE 11.
FT DOMAIN 5429 5466 LAMININ EGF-LIKE 11.
FT DOMAIN 5467 5504 LAMININ EGF-LIKE 11.
FT DOMAIN 5505 5542 LAMININ EGF-LIKE 11.
FT DOMAIN 5543 5580 LAMININ EGF-LIKE 11.
FT DOMAIN 5581 5618 LAMININ EGF-LIKE 11.
FT DOMAIN 5619 5656 LAMININ EGF-LIKE 11.
FT DOMAIN 5657 5694 LAMININ EGF-LIKE 11.
FT DOMAIN 5695 5732 LAMININ EGF-LIKE 11.
FT DOMAIN 5733 5770 LAMININ EGF-LIKE 11.
FT DOMAIN 5771 5808 LAMININ EGF-LIKE 11.
FT DOMAIN 5809 5846 LAMININ EGF-LIKE 11.
FT DOMAIN 5847 5884 LAMININ EGF-LIKE 11.
FT DOMAIN 5885 5922 LAMININ EGF-LIKE 11.
FT DOMAIN 5923 5960 LAMININ EGF-LIKE 11.
FT DOMAIN 5961 5998 LAMININ EGF-LIKE 11.
FT DOMAIN 5999 6036 LAMININ EGF-LIKE 11.
FT DOMAIN 6037 6074 LAMININ EGF-LIKE 11.
FT DOMAIN 6075 6112 LAMININ EGF-LIKE 11.
FT DOMAIN 6113 6150 LAMININ EGF-LIKE 11.
FT DOMAIN 6151 6188 LAMININ EGF-LIKE 11.
FT DOMAIN 6189 6226 LAMININ EGF-LIKE 11.
FT DOMAIN 6227 6264 LAMININ EGF-LIKE 11.
FT DOMAIN 6265 6302 LAMININ EGF-LIKE 11.
FT DOMAIN 6303 6340 LAMININ EGF-LIKE 11.
FT DOMAIN 6341 6378 LAMININ EGF-LIKE 11.
FT DOMAIN 6379 6416 LAMININ EGF-LIKE 11.
FT DOMAIN 6417 6454 LAMININ EGF-LIKE 11.
FT DOMAIN 6455 6492 LAMININ EGF-LIKE 11.
FT DOMAIN 6493 6530 LAMININ EGF-LIKE 11.
FT DOMAIN 6531 6568 LAMININ EGF-LIKE 11.
FT DOMAIN 6569 6606 LAMININ EGF-LIKE 11.
FT DOMAIN 6607 6644 LAMININ EGF-LIKE 11.
FT DOMAIN 6645 6682 LAMININ EGF-LIKE 11.
FT DOMAIN 6683 6720 LAMININ EGF-LIKE 11.
FT DOMAIN 6721 6758 LAMININ EGF-LIKE 11.
FT DOMAIN 6759 6796 LAMININ EGF-LIKE 11.
FT DOMAIN 6797 6834 LAMININ EGF-LIKE 11.
FT DOMAIN 6835 6872 LAMININ EGF-LIKE 11.
FT DOMAIN 6873 6910 LAMININ EGF-LIKE 11.
FT DOMAIN 6911 6948 LAMININ EGF-LIKE 11.
FT DOMAIN 6949 6986 LAMININ EGF-LIKE 11.
FT DOMAIN 6987 7024 LAMININ EGF-LIKE 11.
FT DOMAIN 7025 7062 LAMININ EGF-LIKE 11.
FT DOMAIN 7063 7100 LAMININ EGF-LIKE 11.
FT DOMAIN 7101 7138 LAMININ EGF-LIKE 11.
FT DOMAIN 7139 7176 LAMININ EGF-LIKE 11.
FT DOMAIN 7177 7214 LAMININ EGF-LIKE 11.
FT DOMAIN 7215 7252 LAMININ EGF-LIKE 11.
FT DOMAIN 7253 7290 LAMININ EGF-LIKE 11.
FT DOMAIN 7291 7328 LAMININ EGF-LIKE 11.
FT DOMAIN 7329 7366 LAMININ EGF-LIKE 11.
FT DOMAIN 7367 7404 LAMININ EGF-LIKE 11.
FT DOMAIN 7405 7442 LAMININ EGF-LIKE 11.
FT DOMAIN 7443 7480 LAMININ EGF-LIKE 11.
FT DOMAIN 7481 7518 LAMININ EGF-LIKE 11.
FT DOMAIN 7519 7556 LAMININ EGF-LIKE 11.
FT DOMAIN 7557 7594 LAMININ EGF-LIKE 11.
FT DOMAIN 7595 7632 LAMININ EGF-LIKE 11.
FT DOMAIN 7633 7670 LAMININ EGF-LIKE 11.
FT DOMAIN 7671 7708 LAMININ EGF-LIKE 11.
FT DOMAIN 7709 7746 LAMININ EGF-LIKE 11.
FT DOMAIN 7747 7784 LAMININ EGF-LIKE 11.
FT DOMAIN 7785 7822 LAMININ EGF-LIKE 11.
FT DOMAIN 7823 7860 LAMININ EGF-LIKE 11.
FT DOMAIN 7861 7898 LAMININ EGF-LIKE 11.
FT DOMAIN 7899 7936 LAMININ EGF-LIKE 11.
FT DOMAIN 7937 7974 LAMININ EGF-LIKE 11.
FT DOMAIN 7975 8012 LAMININ EGF-LIKE 11.
FT DOMAIN 8013 8050 LAMININ EGF-LIKE 11.
FT DOMAIN 8051 8088 LAMININ EGF-LIKE 11.
FT DOMAIN 8089 8126 LAMININ EGF-LIKE 11.
FT DOMAIN 8127 8164 LAMININ EGF-LIKE 11.
FT DOMAIN 8165 8202 LAMININ EGF-LIKE 11.
FT DOMAIN 8203 8240 LAMININ EGF-LIKE 11.
FT DOMAIN 8241 8278 LAMININ EGF-LIKE 11.
FT DOMAIN 8279 8316 LAMININ EGF-LIKE 11.
FT DOMAIN 8317 8354 LAMININ EGF-LIKE 11.
FT DOMAIN 8355 8392 LAMININ EGF-LIKE 11.
FT DOMAIN 8393 8430 LAMININ EGF-LIKE 11.
FT DOMAIN 8431 8468 LAMININ EGF-LIKE 11.
FT DOMAIN 8469 8506 LAMININ EGF-LIKE 11.
FT DOMAIN 8507 8544 LAMININ EGF-LIKE 11.
FT DOMAIN 8545 8582 LAMININ EGF-LIKE 11.
FT DOMAIN 8583 8620 LAMININ EGF-LIKE 11.
FT DOMAIN 8621 8658 LAMININ EGF-LIKE 11.
FT DOMAIN 8659 8696 LAMININ EGF-LIKE 11.
FT DOMAIN 8697 8734 LAMININ EGF-LIKE 11.
FT DOMAIN 8735 8772 LAMININ EGF-LIKE 11.
FT DOMAIN 8773 8810 LAMININ EGF-LIKE 11.
FT DOMAIN 8811 8848 LAMININ EGF-LIKE 11.
FT DOMAIN 8849 8886 LAMININ EGF-LIKE 11.
FT DOMAIN 8887 8924 LAMININ EGF-LIKE 11.
FT DOMAIN 8925 8962 LAMININ EGF-LIKE 11.
FT DOMAIN 8963 9000 LAMININ EGF-LIKE 11.
FT DOMAIN 9001 9038 LAMININ EGF-LIKE 11.
FT DOMAIN 9039 9076 LAMININ EGF-LIKE 11.
FT DOMAIN 9077 9114 LAMININ EGF-LIKE 11.
FT DOMAIN 9115 9152 LAMININ EGF-LIKE 11.
FT DOMAIN 9153 9190 LAMININ EGF-LIKE 11.
FT DOMAIN 9191 9228 LAMININ EGF-LIKE 11.
FT DOMAIN 9229 9266 LAMININ EGF-LIKE 11.
FT DOMAIN 9267 9304 LAMININ EGF-LIKE 11.
FT DOMAIN 9305 9342 LAMININ EGF-LIKE 11.
FT DOMAIN 9343 9380 LAMININ EGF-LIKE 11.
FT DOMAIN 9381 9418 LAMININ EGF-LIKE 11.
FT DOMAIN 9419 9456 LAMININ EGF-LIKE 11.
FT DOMAIN 9457 9494 LAMININ EGF-LIKE 11.
FT DOMAIN 9495 9532 LAMININ EGF-LIKE 11.
FT DOMAIN 9533 9570 LAMININ EGF-LIKE 11.
FT DOMAIN 9571 9608 LAMININ EGF-LIKE 11.
FT DOMAIN 9609 9646 LAMININ EGF-LIKE 11.
FT DOMAIN 9647 9684 LAMININ EGF-LIKE 11.
FT DOMAIN 9685 9722 LAMININ EGF-LIKE 11.
FT DOMAIN 9723 9760 LAMININ EGF-LIKE 11.
FT DOMAIN 9761 9798 LAMININ EGF-LIKE 11.
FT DOMAIN 9799 9836 LAMININ EGF-LIKE 11.
FT DOMAIN 9837 9874 LAMININ EGF-LIKE 11.
FT DOMAIN 9875 9912 LAMININ EGF-LIKE 11.
FT DOMAIN 9913 9950 LAMININ EGF-LIKE 11.
FT DOMAIN 9951 9988 LAMININ EGF-LIKE 11.
FT DOMAIN 9989 10000 LAMININ EGF-LIKE 11.

```

Query Match 5.6%, Score 116, DB 1, Length 4393;

Best Local Similarity 24.6%, Pred. No. 0.77; Mismatches 72; Gaps 10;

Matches 55; Conservative 25; Indels 72; Gaps 10;

```

QY 73 YDQGNFTSEMIHNPESDGNIRC-----SLQNSRLHGSAYLVQWMEELFI- 121
DB 3080 -----IGTRPSNHGTYRCASNAKGVQSVNLSVHGPPYVSLPECPVWVK 3127
QY 122 --PSYNL-VVAENPCEVTCPSHMTPLDPI SWEL-----GLVSSSYVYVPESDLOS 173
DB 3128 VGKAVTLECVCAGEPRS-----SAWTRISSTPAKLEQRTYGLMDSH-----T 3170
QY 174 AVSIALTPQSGNLTGCAATKSKARKSATVNLTVIRCPQDPT 217
DB 3171 VLOISSAKPSDAGTYVCLAQNALGTACKQVEVY-----DTG 3207

RESULT 14
NEOL_HUMAN STANDARD; PRT; 1461 AA.
AC Q92859; Q00340;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neogenin precursor.
GN NEOL OR NGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE-Petal brain;
RA Meyerhardt J.A., Look A.T., Bigner S.H., Fearon E.R.;
RT "Identification and characterization of neogenin, a DCC-related
gene."
RL Oncogene 14:1129-1136(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE-Fetal brain;
RA Vielmetter J., Chen X.-N., Miskovich P., Lane R.P., Yamakawa K.,
RT "Molecular characterization of human neogenin, a DCC-related
protein, and the mapping of its gene (NEOL) to chromosomal position 15q22.3-
q23."
RL Genomics 41:414-421(1997).
RN [3]
RP FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: At least 2 ISOFORMS; 1 (SHOWN HERE) AND 2;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED AND ALSO IN CANCER CELL
LINES.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: U61262; AAB17263.1; -.
CC EMBL: U72391; AAC51287.1; -.
CC HSRP: P02751; 17PF.
CC GeneW: HGNC:7754; NEOL.
CC MIM: 601907; -.
CC InterPro: IPR003961; FN_III.

```

DR InterPro: IPR003962; Fail repeat.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; fn3; 6.
 DR PRINTS: PR00014; FNTYEIII.
 DR SMART: SM00060; FN3; 6.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; IgC2; 3.
 KW Transmembrane; Immunoglobulin domain; Glycoprotein; Signal; Repeat;
 KM Alternative splicing
 FT SIGNAL 1 33
 FT CHAIN 34 1461
 FT DOMAIN 34 1461
 FT TRANSMEM 1106 1126
 FT DOMAIN 1127 1461
 FT DOMAIN 67 136
 FT DOMAIN 166 228
 FT DOMAIN 263 327
 FT DOMAIN 353 417
 FT DOMAIN 436 533
 FT DOMAIN 536 629
 FT DOMAIN 630 729
 FT DOMAIN 735 829
 FT DOMAIN 850 950
 FT DOMAIN 951 1052
 FT DOMAIN 1118 1121
 FT DISULFID 74 129
 FT DISULFID 173 221
 FT DISULFID 270 320
 FT CARBOHYD 362 410
 FT CARBOHYD 73 73
 FT CARBOHYD 210 210
 FT CARBOHYD 326 326
 FT CARBOHYD 470 470
 FT CARBOHYD 489 489
 FT CARBOHYD 639 639
 FT CARBOHYD 715 715
 FT CARBOHYD 909 909
 FT VARSPLIC 1248 1300
 FT CONFLICT 168 168
 SQ SEQUENCE 1461 AA; 159958 MW; 7AAE897E69535A21 CRC64;

Query Match 5.6%; Score 114.5; DB 1; Length 1461;
 Best Local Similarity 24.4%; Pred. No. 0.27; Mismatches 107; Indels 73; Gaps 13;
 Matches 68; Conservative 31; Mismatches 107; Indels 73; Gaps 13;

QY 12 SGENETEPONATVYKSGQARENCTVSQGWKLIMMALSDMVLVSRPMEPIITN----- 65
 DB 203 SGMVLT---SNAT---EGDGLRCVYVSGGPP---KTSDEVELKVLDPDEVISDLVFLNQ 254

QY 66 -----DRETSQARD--OGNFTSEMIIHNV 88
 DB 255 PSLPLRVIGDQVILPCVAGSLPPTTIKMKNEALDTESSERIVLLAGS---LEISDV 310

QY 89 EPSDSGNIRCSLONS---LHGSAYLVYQVWGLFIPSVNLYVAENP---PCEVCLPSH 142
 DB 311 TEDAGTYFCIADNGNETIQAQELVQAQPEFLKPTMIYAHESDVIYEEEVGKPT- 369

QY 143 WTRLDPSW-ELGLVSHSYIVPEPSDLSQSAVSLALTPQSNGLTLCVATWKSIAKAR 201
 DB 370 ---PTVAKWKNQDQVIPSDFYKIVEKHNLQ---VLGLVKSDEGEFYQCIADENDVGNQA 421

QY 202 SATVNLTVIRCPDGTGGGINIP-GVLSLPSISGFSLPTW 239
 DB 422 GAOL-ILLEHAPATPTLPSPADRVASIVSTRFKLTW 459

RESULT 15
 RAGE_BOVIN
 ID RAGE_BOVIN STANDARD; PRT; 416 AA.
 AC 028173;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Advanced glycosylation end product-specific receptor precursor
 DE (receptor for advanced glycosylation end products).
 GN AGER OR RAGE.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NX NCBI_TaxId=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Lung;
 RX MEDLINE=92340547; PubMed=1378843;
 RA Nepper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
 RA Elliston K., Stern D., Shaw A.;
 RT Cloning and expression of a cell surface receptor for advanced
 RT glycosylation end products of proteins.
 RL J. Biol. Chem. 267:14998-15004(1992).
 CC - FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 CC RATE IN DIABETES.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC - SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC - SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC - SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M91212; AAA03575.1; -
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; IgC2; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
 FT SIGNAL 1 22
 FT CHAIN 23 416
 FT DOMAIN 23 352
 FT TRANSMEM 353 373
 FT DOMAIN 374 416
 FT DOMAIN 31 105
 FT DOMAIN 136 214
 FT DOMAIN 262 318
 FT DISULFID 38 98
 FT DISULFID 143 207
 FT DISULFID 269 311
 FT CARBOHYD 25 25
 FT CARBOHYD 80 80
 FT DOMAIN 391 396
 SQ SEQUENCE 416 AA; 44182 MW; B703815573E767AE CRC64;

Query Match 5.5%; Score 113.5; DB 1; Length 416;
 Best Local Similarity 23.1%; Pred. No. 0.069; Mismatches 87; Indels 127; Gaps 18;
 Matches 75; Conservative 35; Mismatches 87; Indels 127; Gaps 18;

QY 70 SQRTDGGGFT---SEKITHNVEPSDSGNIR---CS---LONSRLGSAVILVQVGE- 118
 DB 176 TRKHPRTGLTLHSELM---VTPARGALHPTFSCSFPGLRRRALHTAPIDLRVNSEH 232

QY 119 -----LFTPSVNLVVAENP-----CEVTCLEPS---HWTR---LP 147

Wed Apr '30 10:00:20 2003

us-09-729-264-2.rsp

Page 15

```
Db 233 RGGEGPNVDVPLKEVQLV---EPEGAVAPGGTTLTCEADPAQPPQTHWIKDGRPLP 289
QY 148 DISWELGLVSHSSYFYFPEPSDLSQSAVSIALTPOSNGLTLCVATWKSILKARKSATVNL 207
Db 290 -----LPPGMLLPEVG-----PEDGTYSCVATHPSHGPOESRAVS 328
QY 208 TVIRCPD--TGGGINIPGVLSLPSLGSFSLPWGNVGLAGTMLLTPCTLIRCCC 265
Db 329 TITFGEETAGSVEGPG---LETLALTGLG--GLG-----TVALLIGYIVW 373
QY 266 RRRCCGNCNCCCKCRRKRGFRIOFQKKSEKTKNETESGNENSGYNSDEOKTTD 325
Db 374 HRR-----RORNG-----QERKVPENOEERERALENOPE----- 404
QY 326 TASLPPKSCSSDPEORNSSCGPP 349
Db 405 -----EPEAASSTGAP 416
```

Search completed: April 28, 2003, 18:09:37
Job time : 18.2617 secs

•
•
•

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 18:09:54 ; Search time 18.8683 Seconds
(without alignments)
1946.300 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059
Sequence: 1 MGLVFLHSGSGNEVTEGP.....HPQASFTLASPEKVNMTTV 382

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147.5	7.2	404	1	advanced glycosyla
2	133	6.5	1427	2	tumor suppressor -
3	129.5	6.3	164	2	hypothetical prote
4	129.5	6.3	344	2	neurotrophin - rat
5	128	6.2	1447	2	tumor suppressor p
6	127.5	6.2	871	1	protein-tyrosine k
7	127.5	6.2	881	1	protein-tyrosine k
8	125.5	6.1	6642	2	protein UNC-89 - C
9	124	6.0	1070	2	protein-tyrosine k
10	122.5	5.9	188	2	hypothetical prote
11	120	5.8	345	2	oploid-binding pro
12	119	5.8	345	2	oploid-binding pro
13	118	5.7	345	2	oploid-binding pro
14	117.5	5.7	1443	2	oploid-binding pro
15	116.5	5.7	338	2	neogemin - chicken
16	116	5.6	4391	2	perlecan precursor
17	114.5	5.6	152	2	hypothetical prote
18	113.5	5.5	416	2	advanced glycosyla
19	112	5.4	764	2	irregular chiasm C
20	111	5.4	693	2	sodium-dependent p
21	110.5	5.4	620	2	Schwann cell myeli
22	110.5	5.4	1177	2	hypothetical prote
23	109.5	5.3	2491	1	insulin-like growt
24	109	5.3	391	2	butyrophilin homol
25	108.5	5.3	487	2	butyrophilin - mou
26	107	5.2	882	2	receptor tyrosine
27	107	5.2	3707	2	heparan sulfate pr
28	106.5	5.2	662	2	hypothetical prote
29	106.5	5.2	802	2	mitogen and stress

30	106.5	5.2	946	1	A47299	for-related recept
31	106	5.1	1092	1	UN0635	neural cell adhesi
32	104.5	5.1	862	2	I49583	differentiation an
33	104	5.1	2222	2	T13924	sdh protein - fru
34	104	5.1	5825	2	T12117	polyprotein - fava
35	103.5	5.0	890	1	A53743	protein-tyrosine k
36	102.5	5.0	423	2	A31923	amalgam protein pr
37	102.5	5.0	789	2	T29549	hypothetical prote
38	102.5	5.0	1273	2	T12405	hypothetical prote
39	102.5	5.0	1355	2	T28715	sax-3 protein - Ca
40	102.5	5.0	2499	1	A30788	mannose 6-phosphat
41	102.5	5.0	365	2	JC7780	coxsackie- and ade
42	102	5.0	721	2	T41530	hypothetical prote
43	102	5.0	868	2	A46512	CD22 homolog/B-lym
44	102	5.0	880	1	JC4166	protein-tyrosine k
45	102	5.0				

ALIGNMENTS

RESULT 1

161596 advanced glycosylation end-products receptor precursor - human

N:Alternate names: advanced glycosylation end product-binding protein, 35k; glycoprote

C:Species: Homo sapiens (man)

C:Date: 24-May-1996 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999

R:Accession: I61596; B42879; S27968

R:Synonyms: K., 408-419, 1994

A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/

A:Reference number: A55562; M01D:95137587; PMID:7835890

A:Accession: I61596

A:Molecule type: DNA

A:Residues: 1-404 <RES>

A:Cross-references: GB:D28769; NID:9561657; PIDN:BA05958.1; PID:9561659

A:Note: sequence extracted from NCBI backbone (NCBIP:109438)

A:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide

Best Local Similarity 22.6%; Pred. No. 0.0019;
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

QY 15 EVTECPQNAVYKGSQARFNCTVSOG---KLLMALSDMVLSVRMEPIITNDPFTS- 70
Db 125 EIVDSASELTA--GVPKKGVCVSEGSYPAGTSLMHLG-----KPLVNEKGVSV 173
QY 71 ----QRDOGNGFT--SEMIHNVPSDSGNIR-----CSLONSRLHGSAYLVQWGELEF 120
Db 174 KEQTRRHETGTLTLOEELM---VTPARGDPRPTSCSPSPOLPHNRALRTAPIQPRVW 230
QY 121 IP-----SVNLVVAENEP-----CEVTLCPSHWTRLPDLSMELGLVSHSSY 162
Db 231 EPVPLEEVOLV---EPGGAAPAGCTVTLCEVPAQPS-----PQIHMKD----- 274
QY 163 YFVPEPSDLOSANSIILAPPOSNGTLTCVATWKSILKARSAVTNLTVCRODQGGGINTI 222
Db 275 -GVPLEPPSPVILPEISPODGTYSVATSHSGQSRVASTITE--PGEES----- 327
QY 223 PGVLSLPSLGFSLPTWGVGLAGTMLT-----PTCLTRCCCRRCRCGCCGCC 275
Db 328 -----PIAGSVGGSGGLTALALIGLGLTALALIGVILMQR----- 366
QY 276 CRCCFCRRKRRIOPFOKSEKKT--NKETESGENSGNSYSDQKTDLSLPKS 333
Db 367 -----ORGERKAPENOEERAEALN----- 389
QY 334 CESSDEQRNNSCGPP 349
Db 390 -OSEPEAGESSTGPP 404

RESULT 2

151669
tumor suppressor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: 151669
R:Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.
Dev. Biol. 166, 654-665, 1994
A:Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the
A:Reference number: 151668; MID:95113183; PMID:7813784
A:Accession: 151669
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1427 <PIES>
A:Cross-references: EMBL:U10986; NID:9606873; PIDN:AAA70168.1; PID:9606874
C:Genetics:
A:Gene: XDCCA

Query Match 6.5%; Score 133; DB 2; Length 1427;
Best Local Similarity 25.7%; Pred. No. 0.088;
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 17 IEGPQNAVYKGSQARFNCTVSOG---KLLMALSDMVLSVRMEPIITNDPFTS- 74
Db 243 LDRPNSVVAIEQDVLBCAVS-GYPTPIYVMOGD-----EPVIRTR-----KYS 288
QY 75 QGNGFTSEMIHNVPSDSGNIR--SLONSRLHGSAYLVQWGELEFISVNLVVAEN 132
Db 289 YLGG--SMILISNVTDDAGAYTCVATYKKNSTFSADLTVMVPPDLNPAIMLYAESW 346
QY 133 PCEVYCL--PSHWTRLPDISW-ELGLVSHSSYVFPPEPSDLOSASTLALTPDSNGT 187
Db 347 DIEFCAVSGKPS-----PYKWKTKNGEVYIPSDYFQIVDGSNLR-----ILGLVSDSEY 397
QY 188 LTCVA-----TWKSLKARKSATVNLTVI--RCPODTGGGGINIPGVLS 228
Db 398 YOCIAENAGNIQTYADLIIPPAVSSSITLPSAPRDV-----VPVLVSS 442

RESULT 3

T24272

hypothetical protein T01B7.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24272
R:Stims, M.

submitted to the EMBL Data Library, October 1995
A:Reference number: 219867
A:Accession: T24272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <WILD>
A:Cross-references: EMBL:266499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01B7.8
A:Experimental source: clone T01B7
C:Genetics:
A:Gene: CESP:T01B7.8
A:Map position: 2
A:Introns: 20/3; 90/2

Query Match 6.3%; Score 129.5; DB 2; Length 164;
Best Local Similarity 32.1%; Pred. No. 0.015;
Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;

QY 153 LGLVSHSSYVFPPEPSDLOSASTLALTPDSNGTLTCVATWKSILKARSAVTNLTVCIR 212
Db 6 LALLAIGTFTAV---SVOQAV-----LPVSTELATVGTDVASTAIDTLGNSSRV 57
QY 213 PÖDTGGGINIPGVLSLPSLQFSLPTWGVGLAGTMLTPTCTLIRCCCRRCGCC 272
Db 58 KRÖGGCGCCGCC-----GCCGGCGGGG--CGCCCRPRCCCRRCRCC 101
QY 273 --NCCG-RCPCRCR 283
Db 102 CRCCCTRCCTRCR 115

RESULT 4

156551
neurotrophin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 19-May-2000
C:Accession: 156551
R:Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.
J. Neurosci. 15, 2141-2156, 1995
A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed n
A:Reference number: 156551; MID:95198094; PMID:7891157
A:Accession: 156551
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <RES>
A:Cross-references: EMBL:U16845; NID:9755184; PIDN:AAA67445.1; PID:9755185
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amlno-ter

Query Match 6.3%; Score 129.5; DB 2; Length 344;
Best Local Similarity 24.2%; Pred. No. 0.034;
Matches 64; Conservative 43; Mismatches 104; Indels 53; Gaps 14;

QY 4 VIFLHSG-----SGNEVY-EGPQNAVYKGSQARFNCTVSOGKLLMALSDMVLS--- 55
Db 21 LFLVPTGVPYRSGDAPPRKAMDNVTVRQGESATLRCCTIDNRTVRAMLRSTIILAGND 80
QY 56 ---VRMEPIITNDPFTSOGNGFTSEMIHNVPSDSGNIRCSLO-----NSRLH 106
Db 81 KWCIDPRVVLISN---TQTY-----SLETQNDVYDEGPYCSQYTDNHPKTSRVH 129
QY 107 GSATLVQWGELEFISVNLVVAENEPCEVTLPSHWTRLPDISMELGLVSHSSYTVR 166
Db 130 ---LIVQVSPKIVLEISDISINEGNNISLTGATGRPE-PVTVR---HSPKAVGVGS 181
QY 167 EPSDLOSASTLALTPPOSNGTLTCVATWKSILKARSAVTNLTVCIR-----OPTGGGIN 221
Db 182 EDEYLE-----IQGTRGSGETECSSAS-NDVAPAVVRVNTVNVPPYISPAKGTGVPVG 236
QY 222 IPGVL-----SLPSLGFSLPTWKG 241

Db 237 OKT1QCCASAVPSAFPO---WFK 257

RESULT 5

tumor suppressor protein DCC precursor - human
 A:Accession: A54100
 N:Alternate names: colorectal cancer suppressor DCC
 C:Species: Homo sapiens (man)
 C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 05-Nov-1999
 C:Accession: A54100; A40098
 R:Redick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
 Genes Dev. 8, 1174-1183, 1994
 A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.
 A:Reference number: A54100; MUID:95011532; PMID:7926722
 A:Accession: A54100
 A:Molecule type: mRNA
 A:Residues: 1-1447 <HEB>
 A:Cross-references: EMBL:X76132; NID:9453209; PIDN:CAA53735.1; PID:9453210
 R:Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton
 Science 247, 49-56, 1990
 A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.
 A:Reference number: A40098; MUID:90100559; PMID:2294591
 A:Accession: A40098
 A:Molecule type: mRNA
 A:Residues: 1-750 <FEA>
 A:Cross-references: GB:M32292; NID:9181492; PIDN:AAA35751.1; PID:9181493
 C:Genetics:
 A:Gene: GDB:DC
 A:Cross-references: GDB:119838; OMIM:120470
 A:Map position: 18q21.1-18q21.1
 C:Keywords: transmembrane protein; tumor suppressor
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>
 Query Match 6.2%; Score 128; DB 2; Length 1447;
 Best Local Similarity 24.8%; Pred. No. 0.21; Indels 50; Gaps 10;
 Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;
 10 SGGNGV-----IEGPONATVYKSGQARFNCVTSQGW-KLIMWALSMD 51
 Db 220 SRKGNAEVRIISDGLHQLYFLQRPSTVVAIEGKDAVLCCVY-GYPPSPFTWLGE 278
 QY 52 VLVSPMEPIITNDRFTSQRDQGNFSEMTIHNVESDSGNTRC-SLONSRLHSGA 109
 Db 279 VI-----QIRSKYSLLGG--SNLISNVDDDSGMTCTCVYTKNENISASA 323
 QY 110 YLVYVWGLFLTPSVNLVVAENPECEVTCLPSSHMTLPDISW-ELGLVSHSYFVPEP 168
 Db 324 ELTVLPVPPMLNHPNINLXAESMDIEFECTVS-GKPVPTVNMKNDVVIPSDYFOIVGG 382
 QY 169 SDLOSASVILALPPQSNGLTCTVATKSLKARKSATVNLVIRCPDGTGGGINPGVLS 228
 Db 383 SNLR-----ILGVVKSDEGFYQVCAENAGNAQTSQILVFKPAIPSSS-----VLPS 430
 QY 229 LP 230
 Db 431 AP 432
 RESULT 6
 148696
 protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 2 - mouse
 N:Alternate names: receptor-type tyrosine kinase
 N:Contents: protein-tyrosine kinase nsk2 precursor, splice form 4
 C:Species: Mus musculus (house mouse)
 C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
 C:Accession: 148696; S60738
 R:Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D.
 Oncogene 11, 281-290, 1995
 A:Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase
 A:Reference number: 148696; MUID:9534951; PMID:7624144
 A:Accession: 148696

A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-871 <GANI>

A:Cross-references: EMBL:X86444; NID:9929723; PIDN:CAA60165.1; PID:9929724

A:Experimental source: splice form 2

A:Accession: S60738

A:Molecule type: DNA

A:Residues: 1-456; A', 466-871 <GANZ>

A:Cross-references: EMBL:X86444; NID:9929723

A:Experimental source: splice form 4

A:Comment: For alternate splice forms see PIR:148697.

C:Genetics:

A:Gene: nsk2

A:Cross-references: MGI:103308

C:Superfamily: mouse for-related receptor; immunoglobulin homology; protein kinase ho

C:Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; t

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-871/Product: protein-tyrosine kinase nsk2, splice form 2 #status predicted <MAT2

F:22-456; A', 466-871/Product: protein-tyrosine kinase nsk2, splice form 4 #status pre

F:42-101/Domain: immunoglobulin homology <IMM1>

F:135-192/Domain: immunoglobulin homology <IMM2>

F:226-284/Domain: transmembrane #status predicted <TRM>

F:498-518/Domain: transmembrane #status predicted <TRM>

F:575-865/Domain: protein kinase ATP-binding motif

F:583-591/Region: protein kinase ATP-binding motif

F:222,462/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.2%; Score 127.5; DB 1; Length 871;

Best Local Similarity 21.7%; Pred. No. 0.13; Indels 87; Gaps 16;

Matches 70; Conservative 41; Mismatches 125;

QY 15 EYIEGPONATVYKSGQARFNCVTSQGW-KLIMWALSMDVLSRPMPIITNDRFTSQ 73

Db 122 KTRPIPVKIKIEGKAVLPCTMGNRPYSWTKGD-----NALNENRIALE 171

QY 74 DGGNFTSEMTIHNVESDSGNTRC-SLONSRLHSGA-----LVQVMEGLFIPSVNLV 128

Db 172 -----SGSLRHHNVQKEDACQYCVANSLD--GTAAVKLVKEVEVGLILAPESHV 223

QY 129 AENPECVTC-----LPSHMTLPDISW-ELGLVSHSYFVPEPSSDLOSASVILALTPQ 183

Db 224 TFGSEVTLTCTEIGP-----VPTISWLNENNVSSSGISQSVDRYDSRLQLFTKRP 277

QY 184 SNGTTCVAT-----WKSLSKARKSATVNLTVIR-----CPQDTGGGINIPG-----V 225

Db 278 --GLYCIADFNHKGKFSAKAAATVSIAMWSKSDSGYCAQTRGCVLMQGPGEKML 335

QY 226 LSLPSLGSFP-----TWGVYGLAGTMLLPCTTLTRCCCRRCGCCN-- 273

Db 336 LVFLPTSHRDPDAQELLITRAMEL-----KAVSPLCRPAEALICVHLECSFG 388

QY 274 -----CCCRCC-----FCCR 283

Db 389 VVPTPMPICTREYCLAVKELFCAR 411

RESULT 7

148697
 protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 1 - mouse
 N:Alternate names: receptor-type tyrosine kinase
 N:Contents: protein-tyrosine kinase nsk2 precursor, splice form 3
 C:Species: Mus musculus (house mouse)
 C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
 C:Accession: 148697; S60740
 R:Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D.
 Oncogene 11, 281-290, 1995
 A:Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kina
 A:Reference number: 148697; MUID:9534951; PMID:7624144
 A:Accession: 148697
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-881 <GANI>
 A:Cross-references: EMBL:X86445; NID:9929725; PIDN:CAA60166.1; PID:9929726

A: Experimental source: splice form 1
 A: Accession: S60740
 A: Molecule type: DNA
 A: Residues: 1-456, 'A', 466-881 <GAND>
 A: Cross-references: EMBL:X86445; NID:9929725
 A: Experimental source: splice form 3
 C: Comment: For alternate splice forms see PIR:I48696.
 C: Genetics:
 A: Gene: nsk2
 A: Cross-references: MGI:103308
 C: Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase homol
 F: 1-21/Domain: signal sequence #status predicted <Sig>
 F: 22-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <MAT1>
 F: 22-456, 'A', 466-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <MAT1>
 F: 42-101/Domain: immunoglobulin homology <IMM1>
 F: 135-192/Domain: immunoglobulin homology <IMM2>
 F: 226-284/Domain: immunoglobulin homology <IMM3>
 F: 498-518/Domain: transmembrane #status predicted <TRM>
 F: 575-865/Domain: protein kinase homology <KIN>
 F: 583-591/Region: protein kinase AMP-binding motif
 F: 222,462/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match
 Best local similarity 6.2%; Score 127.5; DB 1; Length 881;
 Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

OY 15 EVIEGPNATVLKGSQARFNCVSGWK-LIMWALSDMVLSVRPMEPIITNDFTSQRY 73
 DB 122 KITRPINVKITIEGLKAVLPCTTMNRPVSVWIRGD-----NALRENSRIALF 171
 OY 74 DGGNFTSEMIITHNVEPSDSGNIRCSLONSRLHGSAY-----LTVQWGELEFIPSVNLVY 128
 DB 172 -----SSGLRIHNNQKEDAGQYRCVAKNSL--GTAVSKLVKLEVEVGLIRAPESHNV 223
 OY 129 AENPECEVTC-----LPSHWRLPDISW-ELGLVSHSYFVPEPSDLOSASVILALTPQ 183
 DB 224 TFGSEVTLRCEIGIP-----VPIITWENGNNAVSSGSIQSVKDRVIDRSLFTKRP 277
 OY 184 SNGTLCVAT-----WKSIAKRSATVNLTVIR-----CPQDGGGGINING-----V 225
 DB 278 --GLYTCTATKHKHEKSTAAATVSIAMSKSQKSGCYCAQYNGEVLAMGEGEKML 335
 OY 226 LSLPSLGLFSLP-----TWGKVGGLGLATMLTPTCTITTRCCCCRRRCCGCN 273
 DB 336 LVFLPTTSHRPEDAQGLLHTMANEL-----KAVSPLCRPAEALLCYHLFLECSFG 388
 OY 274 -----CCRCRC-----FCRC 283
 DB 389 VVLPMPICREYGLAVKELEFCAR 411

RESULT 8
 T29757
 Protein UNC-89 - Caenorhabditis elegans
 C: Species: Caenorhabditis elegans
 C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
 C: Accession: T29757
 R: Bu, Z.; Le, T.T.; Wilson, R.
 submitted to the EMBL Data Library, May 1997
 A: Description: The sequence of C. elegans cosmid C09D1.
 A: Reference number: 220679
 A: Accession: T29757
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-6642 <DQZ>
 A: Cross-references: EMBL:AF003131; PTDN:AAB54132.1; GSPDB:GN00019; CESP:unc-89
 A: Experimental source: strain Bristol N2; clone C09D1
 C: Genetics:
 A: Gene: CESP:unc-89
 A: Map position: 1
 A: Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1;
 /3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6603/1

Query Match
 Best local similarity 6.1%; Score 125.5; DB 2; Length 6642;
 Matches 52; Conservative 39; Mismatches 92; Indels 35; Gaps 8;

OY 16 VIEGPNATVLKGSQARFNCVSGWK-LIMWALSDMVLSVRPMEPIITNDFTSQRY 73
 DB 2077 VVDGPKSVTTEKTEFAEFKATIS-GEPAFVTKWTIEKEEESRTITTKTEDVYT----- 2131
 OY 74 DGGNFTSEMIITHNVEPSDSGNIRCSLONSRLHGSAYLTVQWGELEFIPSVNLVY 128
 DB 2132 -----LKSNAKEQGTGVKTAQNASAGDSQADLKPEPVKAKFKRSQJTDKVA 2182
 OY 129 AENPECEVTC-----LPSHWRLPDISWEL-GLVSHSYFVPEPSDLOSASVILALTPQS 184
 DB 2183 DEGEPLKMWNLDELDPGPGT-----EVSWLNGLQPLRSDVYVVDHGDGYHTVIAEKREM 2239
 OY 185 NGTLTCAVATWKSIAKRSATVNLTV-----VIRCPQD 215
 DB 2240 SGTITAKARNNAAGEETSATVYNGNKKPEFVQAQN 2277

RESULT 9
 JC4593
 protein-tyrosine kinase-related receptor PTK7 precursor - human
 N: Alternate names: Receptor protein tyrosine kinase-like protein (RPTK)
 C: Species: Homo sapiens (man)
 C: Date: 16-Apr-1996 #sequence_revision 24-May-1996 #text_change 24-Sep-1999
 C: Accession: JC4593
 R: Park, S.K.; Lee, H.S.; Lee, S.T.
 J. Biochem. 119, 235-239, 1996
 A: Title: Characterization of the human full-length PTK7 cDNA encoding a receptor prot
 A: Reference number: JC4593; MUID:97037064; PMID:8882711
 A: Accession: JC4593
 A: Molecule type: mRNA
 A: Residues: 1-1070 <PAR>
 A: Cross-references: GB:440271; NID:q1322231; PTDN:AAC50484.1; PID:q1322232
 C: Comment: This protein is a member of receptor protein tyrosine kinase family, but, p
 C: Genetics:
 A: Gene: GDB:PTK7
 A: Cross-references: GDB:134760; OMIM:601890
 A: Map position: 6p21.1-6p12.2
 C: Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 F: 1-30/Domain: signal sequence #status predicted <Sig>
 F: 31-1070/Product: cell adhesion; extracellular protein; glycoprotein; phosphotransferase; t
 F: 31-1070/Product: protein-tyrosine kinase 7 #status predicted <EXC>
 F: 704-725/Domain: extracellular #status predicted <EXC>
 F: 726-1070/Domain: transmembrane #status predicted <TRM>
 F: 726-1070/Domain: intracellular #status predicted <INC>
 F: 116,175,184,214,268,283,405,463,567,646/Binding site: carbohydrate (Asn) (covalent)

Query Match
 Best local similarity 6.0%; Score 124; DB 2; Length 1070;
 Matches 58; Conservative 32; Mismatches 103; Indels 36; Gaps 10;

OY 16 VIEGPNATVLKGSQARFNCVSGWK-LIMWALSDMVLSVRPMEPIITNDFTSQRY 74
 DB 227 VVLADQDVVAVYEFEMHCFQSAQPPSLQWLFEDEFTPTNRSRPHLRATVFA----- 282
 OY 75 OGNFTSEMIITHNVEPSDSGNIRCSLONSRLHGSAYLTVQWGELEFIPSVNLVY 129
 DB 283 -----NSLLTLTVRPRNAGIYKICIGQGGKPPITTEALTHLAIEDMDLFEPRVFTAS 337
 OY 130 ENPECEVTCLEPHTRLPDISWEL-GLVSHSYFVPEPSDLOSASVILALTPQSN-G 186
 DB 338 EE---RVTCLEPKGLPPSPVWMEHAGVRLPTEGRVY-----QKHELVLANINAESDAG 387
 OY 187 TLTCVATWKSIAKRSATVNLTV-----IRCPQDGGGINIPGVLSL 229
 DB 388 VYTOHAM-NLAQRQDVAVITATVPSWLKRPQDSLEBGRPGYLDCL 434

RESULT 10

T15651

hypothetical protein C27A2.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15651

R:Nhan, M.

A:Submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid C27A2.

A:Reference number: Z18382

A:Accession: T15651

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-188 <NHA>

A:Cross-references: EMBL:U58760; NID:q1330384; PID:q1330389; PIDN:AA00710.1; GSPDB:GN00

A:Experimental source: strain Bristol N2; clone C27A2

C:Genetics:

A:Gene: CESP:C27A2.5

A:Map position: 2

A:Introns: 19/3; 91/2

Query Match

Best Local Similarity 61.3%; Pred. No. 0.059;

Matches 19; Conservative 1; Mismatches 8; Indels 3; Gaps 2;

OY 256 CTTTTRCCCRRCRCG--NCC-RCCFCCR 283

Db 86 CCCPRKCCCRRCCTCCTCCCTCCTCCTC 116

RESULT 11

JC4025

opioid-binding cell adhesion protein - human

C:Species: Homo sapiens (man)

C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Mar-2000

C:Accession: JC4025

R:Shark, K.B.; Lee, N.M.

Gene 155, 213-317, 1995

A:Title: Cloning, sequencing and localization to chromosome 11 of a cDNA encoding a huma

A:Reference number: JC4025; M0ID:95237612; PMID:7721093

A:Accession: JC4025

A:Molecule type: mRNA

A:Residues: 1-345 <SHA>

A:Cross-references: GB:L34774; NID:q514373; PIDN:AAA36387.1; PID:q514374

A:Experimental source: brain

C:Comment: This protein binds opioid alkaloids in the presence of acidic lipids; exhibit

C:Genetics:

A:Gene: GDB:OPCM; OBCAM; OPCM

A:Cross-references: GDB:251677; OMIM:600632

A:Map position: 11pter-11qter

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match

Best Local Similarity 25.0%; Pred. No. 0.17;

Matches 71; Conservative 37; Mismatches 102; Indels 74; Gaps 18;

OY 4 VIFLHSG---SGNEVI-EGPONATVLKGSQARFNCVSGQWKLIMWALSDMYVL---- 54

Db 21 LFLVPTGVPRSGDATFPKAMDNTVYRQGESATLRCITIDRVTYVAV-LNRSTILVAGN 79

OY 55 ---SVRPMELITNDRTSORYDGGNFTSEMIHNEPSSDGNIRCSLQ-----NSRL 105

Db 80 DKMSIDPRVILLVN---TPQY-----SIMIONVDVDEGPTCSVOYDNDHPTSRV 128

OY 106 HGSATLVYQWGEFLFIPSVNLVVAENEPCEVYTC-----PSHWTLPLDISMELGLVSHSS 161

Db 129 H-----LTYQVPPQIMNISDITVNEGSSVTLCLAIGRPE-----PVTWR-----HLS 173

OY 162 YX-----FVPEPSDLOSASVILATPQSNGLTVCATWKSLEKRSATVNLTV-----IRC 212

Db 174 VKEGGFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRYKVTIVNPPYISK 228

OY 213 PODTGGGINIPGVL-----SLPLSLGFSLPTWGR-----VGLAGLAG 248

Db

229 AKNTGVSQKGLILSCESASVPAEFO---WKEKETRLATGIDG 269

RESULT 12

S03199

opioid-binding protein OPCAM precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 17-Mar-2000

C:Accession: S03199

R:Schofield, P.R.; McFarland, K.C.; Hayflick, J.S.; Wilcox, J.N.; Cho, T.M.; Roy, S.;

EMBO J. 8, 489-495, 1989

A:Title: Molecular characterization of a new immunoglobulin superfamily protein with

A:Reference number: S03199; M0ID:89251576; PMID:2721489

A:Accession: S03199

A:Molecule type: mRNA

A:Residues: 1-345 <SCH>

A:Cross-references: EMBL:X12672; NID:q585; PIDN:CAA31192.1; PID:q586

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-ter

C:Keywords: transmembrane protein

F:1-27/Domain: signal sequence #status predicted <Sig>

F:28-345/Product: opioid-binding protein OPCAM #status predicted <Mat>

Query Match

Best Local Similarity 25.1%; Pred. No. 0.2;

Matches 70; Conservative 36; Mismatches 103; Indels 70; Gaps 17;

OY 4 VIFLHSG---SGNEVI-EGPONATVLKGSQARFNCVSGQWKLIMWALSDMYVL---- 54

Db 21 LFLVPTGVPRSGDATFPKAMDNTVYRQGESATLRCITIDRVTYVAV-LNRSTILVAGN 79

OY 55 ---SVRPMELITNDRTSORYDGGNFTSEMIHNEPSSDGNIRCSLQ-----NSRL 105

Db 80 DKMSIDPRVILLVN---TPQY-----SIMIONVDVDEGPTCSVOYDNDHPTSRV 128

OY 106 HGSATLVYQWGEFLFIPSVNLVVAENEPCEVYTC-----PSHWTLPLDISMELGLVSHSS 161

Db 129 H-----LTYQVPPQIMNISDITVNEGSSVTLCLAIGRPE-----PVTWR-----HLS 173

OY 162 YX-----FVPEPSDLOSASVILATPQSNGLTVCATWKSLEKRSATVNLTV-----IRC 212

Db 174 VKEGGFVSEDEYLE-----ISDIKRDQSGEYECAL-NDVAAPDVRYKVTIVNPPYISK 228

OY 213 PODTGGGINIPGVL-----SLPLSLGFSLPTWGRVGLA 247

Db 229 AKNTGVSQKGLILSCESASVPAEFO---WKEKETRLA 264

RESULT 13

JC1239

opioid-binding protein (clones SG8 and SG13) - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-May-2000

C:Accession: JC1239

R:Lippman, D.A.; Lee, N.M.; Loh, H.H.

Gene 117, 249-254, 1992

A:Title: Opioid-binding cell adhesion molecule (OBCAM)-related clones from a rat bral

A:Reference number: JC1239; M0ID:92347701; PMID:1333369

A:Accession: JC1239

A:Molecule type: mRNA

A:Residues: 1-345 <LIP>

A:Cross-references: GB:M88710; NID:q203247; PIDN:AAA40859.1; PID:q203248; GB:M88711;

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-ter

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 24.7%; Pred. No. 0.24;

Matches 71; Conservative 38; Mismatches 104; Indels 74; Gaps 18;

OY 4 VIFLHSG---SGNEVI-EGPONATVLKGSQARFNCVSGQWKLIMWALSDMYVL---- 54

Db 21 LFLVPTGVPRSGDATFPKAMDNTVYRQGESATLRCITIDRVTYVAV-LNRSTILVAGN 79

QY 55 ---SVRPEPIITNDRTSQRYDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSRL 105
 Db 80 DKMSIDPRVITLVN---TPQY-----SIMIQNVYDDEGPTCSVQTDHNPRTSRV 128
 QY 106 HGSAYLVVQWGELEFISVNIYVAENPECEVTL-----PSHWRLPDISELGLVSHSS 161
 Db 129 H---LIVQVPPQIMNISSDITVNEISSVTLCLAIGRPE-----PVTWR-----HLS 173
 QY 162 YV---FVPEPSDLSQSAVSIALLTPQSGNLTGCVATWKSILKARKSAVNLTV-----IRC 212
 Db 174 VKEGQGVSDDEYLE-----ISDIKROSGEYECGAL-NDVAPDPVRKKTIVNPPYISK 228
 QY 213 PQTGGGINIPGVL---SSLPISGSLPTWGR---VGLAGTML 251
 Db 229 AKNTGVSVGQKGIISCASAVPMAEFQ---WFKEDTRLATGLDGVRI 272

RESULT 14

150600
 neogenin - chicken (fragment)
 C/Species: Gallus gallus (chicken)
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C/Accession: I50600
 R/Velometer: J. Kayem, J.F.; Roman, J.M.; Dreyer, W.J.
 J. Cell Biol. 127, 2005-2020, 1994
 A>Title: Neogenin, an avian cell surface protein expressed during terminal neuronal diff
 A:Reference number: A55193; MUID:95105243; PMID:7806578
 A:Accession: I50600
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1443 <VIP>
 A:Cross-references: EMBL:U07644; NID:9641965; PIDN:AAC59662.1; PID:9641966

Query Match

Best Local Similarity 5.7%; Score 117.5; DB 2; Length 1443;
 Matches 49; Conservative 28; Mismatches 73; Indels 37; Gaps 11;

QY 17 IEQPQATVKGSAQAFNCTVSGW--KLIMALSDMVLVSVRPEPIITNDRTSQRY- 73
 Db 218 VROPSSLTIVTQONAVFPC-VAGGEPPTVVRWTKNG-----EELITED--SERFA 264
 QY 74 -DOGNFTSEMIHNVPSDSGNIRCSLQSAVSIALLTPQSGNLTGCVATWKSILKARKSAVNLTV-----IRC 212
 Db 265 LRAGGS---LLISDYTEEDVGTTCIADNENETEAQALAVQVPEFLKRPANITYAHE 320
 QY 131 NE---PCVATCLPSHWRLPDISW-ELGLVSHSSYFVPEPSDLSQSAVSIALLTPQSN 185
 Db 321 SMDIVFECVETGKPT---PTVKWVKMGDVVIPSDFYKIVKHNMQ---VGLGVKSDE 371
 QY 186 GTLTGVA 192
 Db 372 GFYQCIA 378

RESULT 15

JC1238

opioid-binding protein (clone DU21) - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-May-2000

C/Accession: JC1238

R/Lippman, D.A.; Lee, N.M.; Loh, H.H.

Gene 117, 249-254, 1992

A>Title: Opioid-binding cell adhesion molecule (OBAM)-related clones from a rat brain c

A:Reference number: JC1238; MUID:92347701; PMID:1339369

A:Accession: JC1238

A:Molecule type: mRNA

A:Residues: 1-338 <LIP>

A:Cross-references: GB:M88709; NID:g203245; PIDN:AAA40858.1; PID:g203246

A:Experimental source: brain

C:Genetics:

A:Gene: OBAM

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.7%; Score 116.5; DB 2; Length 338;
 Best Local Similarity 25.0%; Pred. No. 0.31;
 Matches 66; Conservative 33; Mismatches 96; Indels 69; Gaps 16;

QY 22 NATVLKGSQAFNCTVSGWKLIMALSDMVLVSVRPEPIITNDRTSQRYD 74
 Db 37 NNTVQGESATILRCTIIDRVTRVAM-LNRSTILVAGNDKMSIDPRVITLVN---TPQY- 91
 QY 75 QGNFTSEMIHNVPSDSGNIRCSLQ-----NSRLGSAVLTVOVWGELEFISVNLV 128
 Db 92 -----SIMIQNVYDDEGPTCSVQTDHNPRTSRV---LIVQVPPQIMNISSDITV 140
 QY 129 AENPECEVTL-----PSHWRLPDISELGLVSHSSY---FVPEPSDLSQSAVSIALL 180
 Db 141 NEISSVTLCLAIGRPE-----PVTWR-----HLSVKEGQGVSDDEYLE-----ISDI 185
 QY 181 TPQSGNLTGCVATWKSILKARKSAVNLTV-----IRCPDQTGGGINIPGVL---SSLP 231
 Db 186 KRDQSGEYECGAL-NDVAPDPVRKKTIVNPPYISAKRNTGVSVGQKGIISCASAVPM 244
 QY 232 LGFSLPTWGR---VGLAGTML 251
 Db 245 AEFQ---WFKEDTRLATGLDGVRI 265

Search completed: April 28, 2003, 21:08:48
 Job time : 26.8683 secs

Wed Apr 30 10:00:16 2003

us-09-729-264-2.rapb

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 11:05:56 ; Search time 18.8683 Seconds
(without alignments)
1622.277 Million cell updates/sec

Title: US-09-729-264-2
Perfect score: 2059
Sequence: 1 MGLVFLHSGSGNVEIEGP.....HPQAFNLSEKYSNTTVV 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

Rank	Seq ID	Score	Match	Length	DB ID	Description
1	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*	142	6.9	405	1	US-08-755-235-4
2	/cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*	138.5	6.7	2473	9	US-10-184-644-559
3	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*	138.5	6.7	2473	9	US-10-184-634-559
4	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*	136	6.6	708	9	US-10-174-590-584
5	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	136	6.6	708	9	US-10-176-758-584
6	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	136	6.6	708	9	US-10-175-737-584
7	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*	136	6.6	708	9	US-10-173-706-584
8	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*	136	6.6	708	9	US-10-175-738-584
9	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*	136	6.6	708	9	US-10-173-752-584
10	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*	136	6.6	708	9	US-10-176-482-584
11	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*	136	6.6	708	9	US-10-180-552-584
12	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*	136	6.6	708	9	US-10-180-557-584
13	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*	136	6.6	708	9	US-10-173-700-584
14	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*	136	6.6	708	9	US-10-174-572-584

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	142	6.9	405	1	US-08-755-235-4
2	138.5	6.7	2473	9	US-10-184-644-559
3	138.5	6.7	2473	9	US-10-184-634-559
4	136	6.6	708	9	US-10-174-590-584
5	136	6.6	708	9	US-10-176-758-584
6	136	6.6	708	9	US-10-175-737-584
7	136	6.6	708	9	US-10-173-706-584
8	136	6.6	708	9	US-10-175-738-584
9	136	6.6	708	9	US-10-173-752-584
10	136	6.6	708	9	US-10-176-482-584
11	136	6.6	708	9	US-10-180-552-584
12	136	6.6	708	9	US-10-180-557-584
13	136	6.6	708	9	US-10-173-700-584
14	136	6.6	708	9	US-10-174-572-584
15	136	6.6	708	9	US-10-174-579-584
16	136	6.6	708	9	US-10-174-582-584
17	136	6.6	708	9	US-10-174-582-584
18	136	6.6	708	9	US-10-174-582-584
19	136	6.6	708	9	US-10-174-588-584

20	136	6.6	708	9	US-10-175-739-584	Sequence 584, App
21	136	6.6	708	9	US-10-175-740-584	Sequence 584, App
22	136	6.6	708	9	US-10-175-743-584	Sequence 584, App
23	136	6.6	708	9	US-10-176-488-584	Sequence 584, App
24	136	6.6	708	9	US-10-176-492-584	Sequence 584, App
25	136	6.6	708	9	US-10-176-747-584	Sequence 584, App
26	136	6.6	708	9	US-10-176-750-584	Sequence 584, App
27	136	6.6	708	9	US-10-176-985-584	Sequence 584, App
28	136	6.6	708	9	US-10-176-987-584	Sequence 584, App
29	136	6.6	708	9	US-10-176-991-584	Sequence 584, App
30	136	6.6	708	9	US-10-176-992-584	Sequence 584, App
31	136	6.6	708	9	US-10-176-993-584	Sequence 584, App
32	136	6.6	708	9	US-10-184-638-584	Sequence 584, App
33	136	6.6	708	9	US-10-173-695-584	Sequence 584, App
34	136	6.6	708	9	US-10-173-697-584	Sequence 584, App
35	136	6.6	708	9	US-10-173-705-584	Sequence 584, App
36	136	6.6	708	9	US-10-174-576-584	Sequence 584, App
37	136	6.6	708	9	US-10-174-585-584	Sequence 584, App
38	136	6.6	708	9	US-10-174-586-584	Sequence 584, App
39	136	6.6	708	9	US-10-175-747-584	Sequence 584, App
40	136	6.6	708	9	US-10-176-481-584	Sequence 584, App
41	136	6.6	708	9	US-10-176-485-584	Sequence 584, App
42	136	6.6	708	9	US-10-176-487-584	Sequence 584, App
43	136	6.6	708	9	US-10-176-493-584	Sequence 584, App
44	136	6.6	708	9	US-10-176-756-584	Sequence 584, App
45	136	6.6	708	9	US-10-176-911-584	Sequence 584, App

ALIGNMENTS

RESULT 1
US-08-755-235-4
; Sequence 4, Application US/08755235
; Publication No. US20030059423A1
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Wu, Jun
; TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
; FILE REFERENCE: 0575/50159
; CURRENT APPLICATION NUMBER: US/08/755,235
; CURRENT FILING DATE: 1996-11-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Human
US-08-755-235-4

Query Match 6.9%; Score 142; DB 1; Length 405;
Best Local Similarity 22.8%; Pred. No. 0.0048;
Matches 86; Conservative 40; Mismatches 113; Indels 138; Gaps 19;

QY	15	EYIEGPNATVKGSGARENCTVSOG--WKIMALSDWVLSVRPMEDPIITNDRETS	70
DB	125	EYDASSELTA--GVPMKVGTCVSESGSYPACTLSHLDG-----KPLVPMKGVSV	173
QY	71	---QRYDQGNNT--SEMIINVEPSDSGNIR---CSLONSRLHGSATYLVYMGELF	120
DB	174	KEQYRRHPETGELTLOSELTA--VTPARGDDRPTEFSFSGPLRRHRLKTAFLQPRVW	230
QY	121	IP-----SVNLVVAENEP-----CEVYCLPFWWRRLDLSIMELGLVSHSY	162
DB	231	EPVLEEVOLV--EPEGGAVAPGCTVITCEVPAPPS-----PQIHMKD-----	274
QY	163	YFVEPESDASVITLTPQSNCTLTCVATWKSIAKRS--ATVNTVTRCPDPTGGIN	221
DB	275	GVPLPPEPSEVLLPTEIGPDGQGVTSVATHSHGPOESRAVVSITIE--PGEHG----	328
QY	222	IPGVLSLPSLGSFSLPTWKGVGLAGTMTLT-----PTCTVTRCCCRRCGCCNC	274

Db 329 -----PTAGSVGGSGTTLALGILGCTAALLIGVLMOR----- 367
Qy 275 CCRCCFCCRRKRGRIQFOKSEKRT--NKETETESGNGSYNSDEQKTTDTASLPK 332
Db 368 -----ORGBERKAPENOEERERELN----- 390
Qy 333 SCESDPEQRNSCGPP 349
Db 391 ---QSEPEAGESSTGCP 405

RESULT 2

US-10-184-644-559
; Sequence 559, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 559
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-559

Query Match
Best Local Similarity 29.5%; Score 138.5; DB 9; Length 2473;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;
Qy 179 ALTPQSNGLTCAVTWKSLSKRSATVNLVIRCPDGTGGINIRPVLSLSLFSLEPT 238
Db 2274 AATTGAAGTTTCATTAATAATTATATGTTTC----- 2307
Qy 239 GKGVLGAGTMLTPCTLTIRCCCR--RCCGCNCCRCGCC 282
Db 2308 -----ATTCTCATCGCCACCCACCCGCCGCCACCCACC 2344

RESULT 3

US-10-184-634-559
; Sequence 559, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634

; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 559
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-559

Query Match

US-10-174-590-584
; Sequence 584, Application US/10174590
; Publication No. US2003008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-584

RESULT 4

US-10-174-590-584
; Sequence 584, Application US/10174590
; Publication No. US2003008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-584

Query Match
Best Local Similarity 23.8%; Score 136; DB 9; Length 708;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

Qy 3 LVIFLHSGSGNEVEGQNTVTKGSQARENCTVSQWKLIMWALSMDVLSVRMEPI 62
Db 13 LFCFRAGRPSPHFHLDQEDLVLLGGEARLPCLGAYWGLVQMTSGLAGGQR----- 67
Qy 63 ITNDRFTSQRYDQGNFTS--EMIIHNVPSDGSNRCGLNSRLHG--SAVLTV----- 113
Db 68 ---DLPGMSRYWISGNAANGHDLHTRPVELLEBASIEQATQAGLRSPADLVLPPE 124
Qy 114 --QVGELETPSVNLVYAENPCEVTCPSHWTR-LPDISW-ELGLVSHSY--FVPE 167
Db 125 APQVLGS--PVSVLVA--GVPAHLTCRSRGDARPPPELWFRDGVLDGAFHDTLKE 179
Qy 168 --PSDLSAVSILALTPQSNGLTCAVTWKSLSKRSATVNLVIRCPDGT 216
Db 180 GTPGSESTLTLTPFSHDGATFVCRASQALPTGRDIAITLISLQYPEVT 230

RESULT 5

```

LENGTH: 708
TYPE: Prt
ORGANISM: Homo Sapien
US-10-175-737-584

Query Match
Best Local Similarity 23.8%; Pred. No. 0.026;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10.

6.6%; Score 136; DB 9; Length 708;

3 LVILHSGSGNEIECPONATVYKGSQARFNCYTSQGMKTMALSDMVLYSVRMEPI 62
Db 13 LFGRGRAGSPHFLOQPEDVLLGEERLPCALGATWGLQWTKSLAUGOR----- 67
63 ITNDRFTSORDOGNTS---EMIIHNPSPDSGNIRCSLONSJLH-SAYLV----- 113
68 ---DLPGSRKRWISGNMANGOHDHTRPELEDEASLYEQATQACLRSRPAQLHVPRE 124
QY 114 ---QVNGELFDSVNLVVAENPECEVCLPSHWR-LPDISW-ELGLVSHSSYV--FYVE 167
125 APOVYGG---PSVSLVA--GVFANLITRSKGDARPTPELMEFADGVILDGATFHTLKE 179
Db 168 --PSDLOSASILLATPQSNGLTCVAFTWKSLSKARSATVNLVINCQDT 216
QY 168 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::| 230-
Db 180 GPPGSVESTLITPFSHDDGATFVCFARASQALPTGRDTATLTSLOLPEVT

```

RESULT 7
 US-10-173-706-584
 ; Sequence 584, Application US/10173706
 ; Publication No. US2003002293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zhenlin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C7
 ; CURRENT APPLICATION NUMBER: US/10/173,706
 ; CURRENT FILING DATE: 2002-06-17
 ; Prior Application removed - see File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612

[illegible]

Db 180 GTPGSVESTLTLPFSHDDGATFVCARARSQALPTGRDTAITSLOYPEVT 230

RESULT 8

US-10-175-738-584
Sequence 584, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 584
LENGTH: 708
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-584

Query Match
Best Local Similarity 6.6%; Score 136; DB 9; Length 708;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONAVTLKGSQARFNCTVSQGWKLIMALSDMVLSVRPMPI 62
Db 13 LFCFRGRAGSPHFLQDPEDLVLLGEARLPCALGAVGWLTQKSLALGGOR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIIHNEPDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGMSRYWISGMNANGQHDHTRPELEDEASVECCQATQGLRSRPAQLHVLVPE 124
QY 114 --QVMEGLFIPSVNLVVAENEPCEVTCLESHWTR-LPDISW-ELGLVSHSSYY--FYVE 167
Db 125 APQVLGG---PSVSLVA--GVANLTCRSRGDARPTPELLMFRDGVLLDGAFTHQTLKE 179
QY 168 --PSDLOASVSTIALTPQSNCTLCVATWKSIRKRSATVNLTVIRCPDPT 216
Db 180 GTPGSVESTLTLPFSHDDGATFVCARARSQALPTGRDTAITSLOYPEVT 230

RESULT 9

US-10-175-752-584
Sequence 584, Application US/10175752
Publication No. US20030022295A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C60
CURRENT APPLICATION NUMBER: US/10/175,752

;; CURRENT FILING DATE: 2002-06-19
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 584
;; LENGTH: 708
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-175-752-584

Query Match
Best Local Similarity 6.6%; Score 136; DB 9; Length 708;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONAVTLKGSQARFNCTVSQGWKLIMALSDMVLSVRPMPI 62
Db 13 LFCFRGRAGSPHFLQDPEDLVLLGEARLPCALGAVGWLTQKSLALGGOR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIIHNEPDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGMSRYWISGMNANGQHDHTRPELEDEASVECCQATQGLRSRPAQLHVLVPE 124
QY 114 --QVMEGLFIPSVNLVVAENEPCEVTCLESHWTR-LPDISW-ELGLVSHSSYY--FYVE 167
Db 125 APQVLGG---PSVSLVA--GVANLTCRSRGDARPTPELLMFRDGVLLDGAFTHQTLKE 179
QY 168 --PSDLOASVSTIALTPQSNCTLCVATWKSIRKRSATVNLTVIRCPDPT 216
Db 180 GTPGSVESTLTLPFSHDDGATFVCARARSQALPTGRDTAITSLOYPEVT 230

RESULT 10

US-10-176-482-584
Sequence 584, Application US/10176482
Publication No. US20030022296A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C70
CURRENT APPLICATION NUMBER: US/10/176,482
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 584
LENGTH: 708
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-482-584

Query Match
Best Local Similarity 6.6%; Score 136; DB 9; Length 708;
Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONAVTLKGSQARFNCTVSQGWKLIMALSDMVLSVRPMPI 62
Db 13 LFCFRGRAGSPHFLQDPEDLVLLGEARLPCALGAVGWLTQKSLALGGOR----- 67
QY 63 ITNDRFTSQRYDQGNFTS---EMIIHNEPDSGNIRCSLQNSRLHG-SAYLTV----- 113
Db 68 ---DLPGMSRYWISGMNANGQHDHTRPELEDEASVECCQATQGLRSRPAQLHVLVPE 124
QY 114 --QVMEGLFIPSVNLVVAENEPCEVTCLESHWTR-LPDISW-ELGLVSHSSYY--FYVE 167

Db 125 APQVLGG---PSVSLVA--GVPAHLTCRSRGDARPTPELLMFRDGVLLDGAATFHQTLKE 179
 QY 168 --PSDLSAVSIIALTPOSNGLTCVATWKSLSKARSAIVNLTVIRCPDPT 216
 Db 180 GTPGSVESTLTLTPFSHDDGATFVCARSAQLPTGRTAITSLOQYPEVT 230

RESULT 11

US-10-176-757-584
 ; Sequence 584, Application US/10176757
 ; Publication No. US2003002229A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C86
 ; CURRENT APPLICATION NUMBER: US/10/176,757
 ; CURRENT FILING DATE: 2002-06-20
 ; Prior Application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 584
 ; LENGTH: 708
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-176-757-584

Query Match 6.6%; Score 136; DB 9; Length 708;
 Best Local Similarity 23.8%; Pred. No. 0.026;
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVYKGSQARFNCVSGQKLMALSDMVYLSVRPMEPI 62
 Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGIALGGOR----- 67
 QY 63 ITNDRFTSORYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
 Db 68 ---DLPGMSRWISGNANQGHDLHRLPYELDEASVYEQATQAGLRSPQAHLVLPPE 124
 QY 114 --QVMGELFIPSVNLVVAENEPCVTCPLPSHWTR-LPDISW-ELGLLVSHSSY--FVPE 167
 Db 125 APQVLGG---PSVSLVA--GVPAHLTCRSRGDARPTPELLMFRDGVLLDGAATFHQTLKE 179
 QY 168 --PSDLSAVSIIALTPOSNGLTCVATWKSLSKARSAIVNLTVIRCPDPT 216
 Db 180 GTPGSVESTLTLTPFSHDDGATFVCARSAQLPTGRTAITSLOQYPEVT 230

RESULT 12

US-10-176-913-584
 ; Sequence 584, Application US/10176913
 ; Publication No. US20030022298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ;; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ;; FILE REFERENCE: P3430R1C66
 ;; CURRENT APPLICATION NUMBER: US/10/176,913
 ;; CURRENT FILING DATE: 2002-06-20
 ;; Prior Application removed - See file wrapper or Palm
 ;; NUMBER OF SEQ ID NOS: 612
 ;; SEQ ID NO 584
 ;; LENGTH: 708
 ;; TYPE: PRT
 ;; ORGANISM: Homo Sapien
 ; US-10-176-913-584

Query Match 6.6%; Score 136; DB 9; Length 708;
 Best Local Similarity 23.8%; Pred. No. 0.026;
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVYKGSQARFNCVSGQKLMALSDMVYLSVRPMEPI 62
 Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGIALGGOR----- 67
 QY 63 ITNDRFTSORYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
 Db 68 ---DLPGMSRWISGNANQGHDLHRLPYELDEASVYEQATQAGLRSPQAHLVLPPE 124
 QY 114 --QVMGELFIPSVNLVVAENEPCVTCPLPSHWTR-LPDISW-ELGLLVSHSSY--FVPE 167
 Db 125 APQVLGG---PSVSLVA--GVPAHLTCRSRGDARPTPELLMFRDGVLLDGAATFHQTLKE 179
 QY 168 --PSDLSAVSIIALTPOSNGLTCVATWKSLSKARSAIVNLTVIRCPDPT 216
 Db 180 GTPGSVESTLTLTPFSHDDGATFVCARSAQLPTGRTAITSLOQYPEVT 230

RESULT 13

US-10-180-552-584
 ; Sequence 584, Application US/10180552
 ; Publication No. US20030022300A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C153
 ; CURRENT APPLICATION NUMBER: US/10/180,552
 ; CURRENT FILING DATE: 2002-06-25
 ; Prior Application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 584
 ; LENGTH: 708
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-180-552-584

Query Match 6.6%; Score 136; DB 9; Length 708;
 Best Local Similarity 23.8%; Pred. No. 0.026;
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;

QY 3 LVIFLHSGSGNEVIEGPONATVYKGSQARFNCVSGQKLMALSDMVYLSVRPMEPI 62
 Db 13 LFCFRGRAGSPHFLQOPEDLVLLGGEARLPCALGAYWGLVQWTKSGIALGGOR----- 67
 QY 63 ITNDRFTSORYDQGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 18:14:43 ; Search time 13.9029 Seconds
(without alignments)
808.430 Million cell updates/sec

Title: US-09-729-264-2

Perfect score: 2059
Sequence: 1 MGLVIFLHGSGSGNEVIEGP.....HPQASFLASPEKYSNTTV 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/5C.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/5D.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/5E.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/5F.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128.5	6.2	1345	2	US-08-977-767-3
2	128	6.2	1447	4	US-09-041-886-25
3	128	6.2	1447	5	PCT-US94-05277-2
4	126.5	6.1	869	1	US-08-374-834-16
5	126.5	6.1	869	2	US-08-644-271-29
6	126.5	6.1	869	4	US-09-077-955-33
7	123.5	6.0	332	4	US-09-062-365-1
8	122.5	5.9	340	4	US-09-651-200-2
9	122.5	5.9	441	4	US-09-651-200-6
10	121.5	5.9	534	4	US-09-651-200-24
11	121.5	5.9	534	2	US-08-633-148-4
12	118.5	5.8	340	2	US-08-633-148-2
13	118.5	5.8	340	2	US-08-633-148-2
14	117	5.7	1395	4	US-08-540-245A-15
15	114	5.5	868	2	US-08-374-834-1
16	112	5.4	868	2	US-08-644-271-1
17	112	5.4	868	4	US-09-077-955-1
18	112	5.4	868	4	US-09-499-964-1
19	110	5.3	478	5	PCT-US95-08493-15
20	109.5	5.3	860	5	PCT-US95-08493-19
21	109.5	5.3	868	5	PCT-US95-08493-21
22	107.5	5.2	362	1	US-08-415-751-6
23	107.5	5.2	362	1	US-08-979-424-3
24	107	5.2	365	4	US-09-272-436-2
25	106.5	5.2	946	5	PCT-US95-08493-13
26	106	5.1	319	1	US-08-597-495B-22
27					

28	106	5.1	319	4	US-09-068-051A-22	Sequence 22, Appl
29	106	5.1	319	4	US-09-336-536-67	Sequence 67, Appl
30	106	5.1	319	4	US-09-254-465A-6	Sequence 6, Appl
31	106	5.1	801	1	US-07-906-349A-6	Sequence 6, Appl
32	105	5.1	285	3	US-08-482-085B-20	Sequence 20, Appl
33	105	5.1	1297	4	US-09-540-245A-17	Sequence 17, Appl
34	105	5.1	1381	4	US-09-540-245A-16	Sequence 16, Appl
35	104.5	5.1	879	1	US-08-554-612C-1	Sequence 1, Appl
36	104	5.1	421	2	US-08-659-984A-1	Sequence 1, Appl
37	104	5.1	421	2	US-08-659-984A-1	Sequence 1, Appl
38	104	5.1	444	2	US-08-659-984A-5	Sequence 5, Appl
39	104	5.1	444	4	US-08-660-331-5	Sequence 2, Appl
40	103.5	5.0	890	1	US-08-445-640-2	Sequence 2, Appl
41	103.5	5.0	890	3	US-08-170-558-2	Sequence 2, Appl
42	103.5	5.0	890	3	US-08-447-314-2	Sequence 2, Appl
43	103.5	5.0	890	3	US-08-445-461-2	Sequence 2, Appl
44	103.5	5.0	911	1	US-08-286-305A-1	Sequence 1, Appl
45	103.5	5.0	911	2	US-08-441-104A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-977-767-3
Sequence 3, Application US/08977767

Patent No. 5972684

GENERAL INFORMATION:
APPLICANT: Bandman, Olga

APPLICANT: Yee, Henry

APPLICANT: Greenwald, Sara

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: CARBONIC ANHYDRASE VIII

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/977,767

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0423 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1345 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1532042

US-08-977-767-3

Query Match 6.2% Score 128.5; DB 2; Length 1345;

Best Local Similarity 33.0%; Pred. No. 0.0095;
Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5;

QY 186 GTTCVATWKLKARKSATVNLVIRCPDGTGGT-----NIPGVLSLPSLPSLPTWKG 241
Db 414 GTCTCTGT-----GC-CGTGGAGCGTCAGAGCCCGCGTGGATGTGA 455
QY 242 VGLGACTMTLT-PTCTLTTRCCCRRCGCCGCCRC-----CFCC 282
Db 456 CGTGAGAGGGCTCTCATGACCCCTTCCGTGCCCTCTGAGACGTACGCCAC 507

RESULT 2

US-09-041-886-25
Sequence 25, Application US/09041886
Patent No. 6235872

GENERAL INFORMATION:

APPLICANT: Bredesen, Dale E.
APPLICANT: Rabizadeh, Sharroz
TITLE OF INVENTION: Proapoptotic peptides, Dependence
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESSES:
ADDRESSER: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-886-25

Query Match
Best Local Similarity 24.8%; Score 128; DB 4; Length 1447;
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 10 SGSGNEV-----IEGPNATVLRKSGARFNCVSGW-KLIMNALSDM 51
Db 220 SRTGNEAEVRLSDPGLHRLQYFLQRPNSVVAIEGDAVLECCVS-GYPPPSFTWLRGE 278
QY 52 VLVSVRMEPIITNDRTSQRYDOGGNFTSEMIINHVESDSGNIRC--SIQNSRLHGA 109
Db 279 VI-----QIRSKKYSLLGG--SNLISNVTDGSGMYTCVYTKNENISASA 323
QY 110 YLVQVAGELFISVNLVVAENEPCEVTCIPSHWTRLPDISW-ELGLVSHSSYFVPEP 168
Db 324 ELFLVLPPEWFLNPSNLVYAESMDIEFFECTVS-GKPVPTVNMKMGDVVIRPSDYFQIVG 382
QY 169 SDIQASVSIATLPQSGNGLTCVATWKLKARKSATVNLVIRCPDGTGGGGINIPGVLS 228
Db 383 SNLR-----ILGVKSDGEFYQCAVAENEGMAGTSAQLIVPRPAIPSSS-----VLPS 430

QY 229 LP 230
Db 431 AP 432

RESULT 3

PCT-US94-05277-2
Sequence 2, Application PC/TUS9405277
GENERAL INFORMATION:

APPLICANT: Bruskin, Arthur
APPLICANT: Jarosz, David E.
APPLICANT: Johnson, Karen
APPLICANT: Kinzier, Kenneth W.
APPLICANT: Vogelstein, Bert
APPLICANT: Zabrecky, James R.
TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSER: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05277
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42709
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05277-2

Query Match
Best Local Similarity 6.2%; Score 128; DB 5; Length 1447;
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 10 SGSGNEV-----IEGPNATVLRKSGARFNCVSGW-KLIMNALSDM 51
Db 220 SRTGNEAEVRLSDPGLHRLQYFLQRPNSVVAIEGDAVLECCVS-GYPPPSFTWLRGE 278
QY 52 VLVSVRMEPIITNDRTSQRYDOGGNFTSEMIINHVESDSGNIRC--SIQNSRLHGA 109
Db 279 VI-----QIRSKKYSLLGG--SNLISNVTDGSGMYTCVYTKNENISASA 323
QY 110 YLVQVAGELFISVNLVVAENEPCEVTCIPSHWTRLPDISW-ELGLVSHSSYFVPEP 168
Db 324 ELFLVLPPEWFLNPSNLVYAESMDIEFFECTVS-GKPVPTVNMKMGDVVIRPSDYFQIVG 382
QY 169 SDIQASVSIATLPQSGNGLTCVATWKLKARKSATVNLVIRCPDGTGGGGINIPGVLS 228
Db 383 SNLR-----ILGVKSDGEFYQCAVAENEGMAGTSAQLIVPRPAIPSSS-----VLPS 430
QY 229 LP 230
Db 431 AP 432

RESULT 4
US-08-374-834-16
Sequence 16, Application US/08374834
Patent No. 5656473

GENERAL INFORMATION:

APPLICANT: Valenzuela, et al.
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,834
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/095,658
FILING DATE: 21-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J.
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 190A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 345-7400
TELEFAX: (914) 345-7721

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:
LENGTH: 869 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-374-834-16

Query Match 6.1%; Score 126.5; DB 1; Length 869;

Best Local Similarity 20.9%; Pred. No. 0.0078;
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

QY 15 EVIEGPNATVLKGSQARENCYVSQGMK-LIMWALSDMVVLSVRPMEPIITNDRTSORY 73
DB 122 KITRPPIINVKIIEGKRAVLPCITMGNPKPSVSIKGD-----SPLRENSRLAVLE- 171
QY 74 DQGNFTSEMIINWEPDSGNIRCSLQNSRLHGSAY-LTYQVMEGLFIPSYNLVVAENE 132
DB 172 -GSLRIHNVQKEDAGQRCVAKNSL--GTAISKYVKLEVEVFARILRAPESHV 223
QY 133 P-----CEVCLPSSHMTRLPDISW-ELGLVSHSSYYFVPEPSDLOSANSILALTPQ 183
DB 224 TFGSFVTLHCTATGIP-----VPIITWIENGNAVSSGSIQESYKDVYDSRLQLFTTKP- 277
QY 184 SNGTLTCVAT-----WKSIAKRSATVNLTYIRCDPTGG-----INIGVLSLP 230
DB 278 --GLTTCIATNMKGKFEKSTAKAATISIAEWSKPKQKNGKCAQYRGEVCNAYLADALV 335
QY 231 SLGFSL-----PTWGKVGGLAGTMLTPTCTLITRCCCRRCOCGN----- 273
DB 336 FLNTSYADPEEAQELLVHTANML-----KVSPYCPRAEALLCNHIFQECSPGVVP 388
QY 274 -----CCCRCC-----FCCRRKRGFRIOFKKSEKKTNETETESNGNSGVNSDQKT 323
DB 389 TPLPICREGLAVKELFCAKE---WLVMEKETHRGILYRSEMHLLSYPECKLPSMMDP 444
QY 324 TDTASLPPKSCSSDPFORNSSCGPPIHQADQRP-----PRASHPOASFNLA 371

DB 445 TACARLP-----HDIYNKENTKTEPP--MTSSKPSVDIPNLPSSSSSSESVS 489

RESULT 5

US-08-644-271-29

Sequence 29, Application US/08644271

Patent No. 5814478

GENERAL INFORMATION:

APPLICANT: Valenzuela, et al.
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/644,271
FILING DATE: 10-MAY-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/008,657
FILING DATE: 15-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J.
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 195A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:
LENGTH: 869 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-644-271-29

Query Match 6.1%; Score 126.5; DB 2; Length 869;

Best Local Similarity 20.9%; Pred. No. 0.0078;
Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

QY 15 EVIEGPNATVLKGSQARENCYVSQGMK-LIMWALSDMVVLSVRPMEPIITNDRTSORY 73
DB 122 KITRPPIINVKIIEGKRAVLPCITMGNPKPSVSIKGD-----SPLRENSRLAVLE- 171
QY 74 DQGNFTSEMIINWEPDSGNIRCSLQNSRLHGSAY-LTYQVMEGLFIPSYNLVVAENE 132
DB 172 -GSLRIHNVQKEDAGQRCVAKNSL--GTAISKYVKLEVEVFARILRAPESHV 223
QY 133 P-----CEVCLPSSHMTRLPDISW-ELGLVSHSSYYFVPEPSDLOSANSILALTPQ 183
DB 224 TFGSFVTLHCTATGIP-----VPIITWIENGNAVSSGSIQESYKDVYDSRLQLFTTKP- 277
QY 184 SNGTLTCVAT-----WKSIAKRSATVNLTYIRCDPTGG-----INIGVLSLP 230
DB 278 --GLTTCIATNMKGKFEKSTAKAATISIAEWSKPKQKNGKCAQYRGEVCNAYLADALV 335
QY 231 SLGFSL-----PTWGKVGGLAGTMLTPTCTLITRCCCRRCOCGN----- 273
DB 336 FLNTSYADPEEAQELLVHTANML-----KVSPYCPRAEALLCNHIFQECSPGVVP 388

OY 274 -----CCRCRC-----FCRRKRGFRLOPKSEKTEKNETTESGNGSNGSDEOKT 323
 Db 389 TPICREYCLAVKELCAKE-----WLVMEKTHRGLYSEMHLLSVPCSKLPSMHMDP 444
 OY 324 TDTASLPKSCGESSDPQRNNSCGPPHQRADOP-----PRPASHPOASFLA 371
 Db 445 TACARLP-----HLDYKNEMLKTFPP--MTSSKPSVDIDINLPSSSSSSPTSVS 489

RESULT 6

US-09-077-955-33
 ; Sequence 33, Application US/09077955A
 ; Patent No. 6413740
 ; GENERAL INFORMATION:

APPLICANT: Valenzuela et al., David M.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
 FILE REFERENCE: REG195-B-PCT-US
 CURRENT APPLICATION NUMBER: US/09/077,955A
 EARLIER FILING DATE: 1998-09-10
 EARLIER APPLICATION NUMBER: PCT/US96/20696
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 08/644,271
 EARLIER FILING DATE: 1996-05-10
 EARLIER APPLICATION NUMBER: 60/008,657
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 33
 LENGTH: 869
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-077-955-33

Query Match
 Best Local Similarity 6.1%; Score 126.5; DB 4; Length 869;
 Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

OY 15 EVIEGPONATVYLGSOARENCTVSOGWK-LIMWALSDMVVLSVRPMEPIITNDRTSQRY 73
 Db 122 KITRPPINVKRIIEGLKAVLPCTMGMPKPSVSIKGD-----SPRENSRLAVLE- 171
 OY 74 DGGNFTSEMIINHPSPSGNTRCSLONSRLHGSNY-LTYQWMEGLFIPSVNLVAENE 132
 Db 172 -----SGLRIHNVOKEDAGQRCYAKNSL--GTAYSKRVKLEVEVFAIRLIRAPESHNV 223
 OY 133 P-----CEYCLPSSHTRLPDISW-ELGLVSHSSVYFVPEPSDQSAVSLALTPQ 183
 Db 224 TRESVTLHCTATGIP-----VPTTWIENGNAVSSGSIQESVDRVIDSLQLEFTRP- 277
 OY 184 SNGTLVCVAT-----WKSILKARKSATVNLVIRCPQDTGGC-----INIPGVLSSLP 230
 Db 278 --GLYTCIATNKHGKFTAKRAAATISIAEWSKPOKDKGCAQYRGECVNAVLAKDALV 335
 OY 231 SLGFSL-----PTWKGVLGLAGTMLLPTCTLLTRCCCRRCGCCGN----- 273
 Db 336 FLNTSYADPEEAQELLVTANNEL-----KVYSPVCRAAEALMLLCNHLFQECSEGVVP 388
 OY 274 -----CCRCRC-----FCRRKRGFRLOPKSEKTEKNETTESGNGSNGSDEOKT 323
 Db 389 TPICREYCLAVKELCAKE-----WLVMEKTHRGLYSEMHLLSVPCSKLPSMHMDP 444
 OY 324 TDTASLPKSCGESSDPQRNNSCGPPHQRADOP-----PRPASHPOASFLA 371
 Db 445 TACARLP-----HLDYKNEMLKTFPP--MTSSKPSVDIDINLPSSSSSSPTSVS 489

RESULT 7

US-09-062-365-1
 ; Sequence 1, Application US/09062365
 ; Patent No. 6465422
 ; GENERAL INFORMATION:
 ; APPLICANT: Schmidt, Ann Marie
 ; APPLICANT: Stern, David

;; TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
 ;; FILE REFERENCE: SUBJECT
 ;; CURRENT APPLICATION NUMBER: US/09/062,365
 ;; CURRENT FILING DATE: 1998-04-17
 ;; NUMBER OF SEQ ID NOS: 6
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 1
 ;; LENGTH: 332
 ;; TYPE: PRT
 ;; ORGANISM: Human
 US-09-062-365-1

Query Match
 Best Local Similarity 6.0%; Score 123.5; DB 4; Length 332;
 Matches 65; Conservative 33; Mismatches 89; Indels 83; Gaps 14;

OY 15 EVIEGPONATVYLGSOARENCTVSOG--KLIWALSDMVVLSVRPMEPIITNDRTSQRY- 70
 Db 103 EIVDSASELTA--GVPMKVGTCVSEGSTPAGTYSWHIDG-----KPLVNEKGVSV 151
 OY 71 -----ORVGGNFT--SEMIHNVPEPSDGNIR-----CSLONSRLHGSAYLTYQWMEGLF 120
 Db 152 KEOTRHPETGELFTLOSELM--VTPARGDDPRPTFSCSPCLPRHRLKTRAPLOPRW 208
 OY 121 IP-----SVNLVAENEP-----CEYCLPSSHTRLPDISWELGLLVSHSSY 162
 Db 209 EPVPLEEYOLV--EPEGGAVAPGCVTLTCEVPAQPS-----PQIHMKD----- 252
 OY 163 YFVPEPSDLOSANSIILATFQNSGTLTCVATWKSILKARKSATVNLVIRCPQDTGGGGINI 222
 Db 253 -GVPLPPEPSPVILPEIGPODGYSCVATSHSHQPSQPSRAVSIISIE-POEEC----- 305
 OY 223 PGVLSLPSLGFSLPWKGVLGLACTMVL 252
 Db 306 -----PTAGSVGGSGGLCTLAL 321

RESULT 8

US-09-651-200-2
 ; Sequence 2, Application US/09651200
 ; Patent No. 6429303
 ; GENERAL INFORMATION:
 ; APPLICANT: Green et al
 ; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 ; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 ; FILE REFERENCE: 15966-362 (CURA-62)
 ; CURRENT APPLICATION NUMBER: US/09/651,200
 ; CURRENT FILING DATE: 2000-08-30
 ; PRIOR APPLICATION NUMBER: 60/152383
 ; PRIOR FILING DATE: 1999-09-03
 ; PRIOR APPLICATION NUMBER: 60/172909
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/183578
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 340
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-651-200-2

Query Match
 Best Local Similarity 5.9%; Score 122.5; DB 4; Length 340;
 Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;

OY 10 SGGNEVIEGPNATVYLGSOARENCTVS--QGWKL-----IMWALSDMVVLSVRPMEPI 62
 Db 48 SPTGAVEVQVPEPVALVATGATLHCSFSPFSLTQULWLIQULDTDTQGLV----- 100
 OY 63 ITNDRTSQRYDOGGNF-----TSEMIHNVPEPSDGNIRCSLONSRLHGS 108

```

Db 101 ---HSFTBGR-DGSAVANRTALEPDLAOGNASTLRQRYVADSGSFTCFV-SINDFS 155
QY 109 AYLTVQVNGELFIPSVNLV-----VAENPECEVTCIPSHWTRLP--DISWELG---LYSH 159
Db 156 AAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPAEVFWODGQVPLTGN 213
QY 160 SSIYFVPEPSDQSAVSIALTPQSNGLTCVATWKSILKARKSATVNLTVIRCP---QDT 216
Db 214 VTTSQMANEQGLFDVHSVLRVVLGANGTYS-----LVARNVLOODA 255
QY 217 GGGINIPGVLSLPSLGSFLPTWKGVLGAGTMLTPCTLTIRCCCCRRCCGCCNCCC 276
Db 256 HGSVTTTGQPMTPPEAL-----WTVGLSVCLIALV-----288
QY 277 RCFCCRRKRGFRIOFOKSEKTKETETESGENSGYNSDEOKTDTASLPKSCS 336
Db 289 ALAFVCMRK-----IKOSCEENNAAGDODG-----EGGSKTALQPLKHS 331
QY 337 SDPE 340
Db 332 KEDD 335

```

RESULT 9
US-09-651-200-4
Sequence 4, Application US/09651200
Patent No. 6429303

```

GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 4
LENGTH: 441
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-200-4

```

Query Match 5.9%; Score 122.5; DB 4; Length 441;
Best local similarity 21.2%; Pred. No. 0.007; Indels 109; Gaps 17;
Matches 77; Conservative 50; Mismatches 128;

```

QY 10 SSGSNEVIEGPNATV-LKGSQARFNCYS--QGKTL-----IMWALSMDVNLVSVRPEPI 62
Db 149 SPGAEVQVPEDPVALVGTDTATLRCSPSPGFSLTQNLIMQDLTDKQLV-----201
QY 63 ITNDRFTSQRYDQGNF-----TSEMIIHNVPEPSDSGNIRSLQNSRLHGS 108
Db 202 ---HSFTBGR-DGSAVANRTALEPDLAOGNASTLRQRYVADSGSFTCFV-SINDFS 256
QY 109 AYLTVQVNGELFIPSVNLV-----VAENPECEVTCIPSHWTRLP--DISWELG---LYSH 159
Db 257 AAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPAEVFWODGQVPLTGN 314
QY 160 SSIYFVPEPSDQSAVSIALTPQSNGLTCVATWKSILKARKSATVNLTVIRCP---QDT 216
Db 315 VTTSQMANEQGLFDVHSVLRVVLGANGTYS-----LVARNVLOODA 356
QY 217 GGGINIPGVLSLPSLGSFLPTWKGVLGAGTMLTPCTLTIRCCCCRRCCGCCNCCC 276
Db 357 HGSVTTTGQPMTPPEAL-----WTVGLSVCLIALV-----389

```

```

QY 277 RCFCCRRKRGFRIOFOKSEKTKETETESGENSGYNSDEOKTDTASLPKSCS 336
Db 390 ALAFVCMRK-----IKOSCEENNAAGDODG-----EGGSKTALQPLKHS 432
QY 337 SDPE 340
Db 433 KEDD 436

```

RESULT 10
US-09-651-200-6
Sequence 6, Application US/09651200
Patent No. 6429303

```

GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 6
LENGTH: 534
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-200-6

```

Query Match 5.9%; Score 121.5; DB 4; Length 534;
Best local similarity 21.2%; Pred. No. 0.011; Indels 109; Gaps 17;
Matches 77; Conservative 49; Mismatches 129;

```

QY 10 SSGSNEVIEGPNATV-LKGSQARFNCYS--QGKTL-----IMWALSMDVNLVSVRPEPI 62
Db 242 SPGAEVQVPEDPVALVGTDTATLRCSPSPGFSLTQNLIMQDLTDKQLV-----294
QY 63 ITNDRFTSQRYDQGNF-----TSEMIIHNVPEPSDSGNIRSLQNSRLHGS 108
Db 295 ---HSFTBGR-DGSAVANRTALEPDLAOGNASTLRQRYVADSGSFTCFV-SINDFS 349
QY 109 AYLTVQVNGELFIPSVNLV-----VAENPECEVTCIPSHWTRLP--DISWELG---LYSH 159
Db 350 AAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITC--SSYRGYPAEVFWODGQVPLTGN 407
QY 160 SSIYFVPEPSDQSAVSIALTPQSNGLTCVATWKSILKARKSATVNLTVIRCP---QDT 216
Db 408 VTTSQMANEQGLFDVHSVLRVVLGANGTYS-----LVARNVLOODA 449
QY 217 GGGINIPGVLSLPSLGSFLPTWKGVLGAGTMLTPCTLTIRCCCCRRCCGCCNCCC 276
Db 450 HGSVTTTGQPMTPPEAL-----WTVGLSVCLIALV-----482
QY 277 RCFCCRRKRGFRIOFOKSEKTKETETESGENSGYNSDEOKTDTASLPKSCS 336
Db 483 ALAFVCMRK-----IKOSCEENNAAGDODG-----EGGSKTALQPLKHS 525
QY 337 SDPE 340
Db 526 KEDD 529

```

RESULT 11
US-09-651-200-24
Sequence 24, Application US/09651200
Patent No. 6429303

GENERAL INFORMATION:
 APPLICANT: Green et al
 TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 FILE REFERENCE: 15966-562 (CURA-62)
 CURRENT APPLICATION NUMBER: 05/09/651,200
 PRIOR FILING DATE: 2000-08-30
 PRIOR APPLICATION NUMBER: 60/152383
 PRIOR FILING DATE: 1999-09-03
 PRIOR APPLICATION NUMBER: 60/172909
 PRIOR FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: 60/183578
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 24
 LENGTH: 534
 TYPE: PRT
 ORGANISM: Unknown
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: Sequence
 US-09-651-200-24

Query Match
 Best Local Similarity 5.9%; Score 121.5; DB 4; Length 534;
 Matches 77; Conservative 49; Mismatches 129; Indels 109; Gaps 17;

QY 10 SGSGNEVEGPNATV-LKSGQARFNCTVS--QGWKL---IMWALSDMVVLSVRMEPI 62
 DB 242 SPTGAVEQVPEDEYVALVGDVTLRCSPSPGFSFLAOLNLTWOLTRKOLV----- 294
 QY 63 ITNREFSORYDOGNF-----TSEMIIHNVPEPDSGNINCSLQNSRLHS 108
 DB 295 ---HSFTEGR-DOGSAYANTALFPPDLAQGNASLRQVRADSEFTCEV-STIRDS 349
 QY 109 AYLTVQVAGELFIPSVLV---VAENPERCVTCLPSHMTLP--DISWELGL---LVSH 159
 DB 350 AAVSLQVAARYSPKMTLPENKDLRGDVTYITC--SSYRGVPEAEVFDQGVPLIGN 407
 QY 160 SYTFVEPEPDLQSAVSIATLPQSNGLTCVATWKSILKARKSATVNLVIRCP---QDT 216
 DB 408 VTTQSQANDEQGLFDVHSHVRLVGLANGTYSC-----LVANPVLQDDA 449
 QY 217 GGGINIPGVLSLPSLGFSLPTWKGVLGLAGTMTLLPTCTLTTRCCRRRCGCCNCCC 276
 DB 450 HGSVTLTGQPMTFPEAL---WTVGLSVCLIALLV----- 482
 QY 277 KCFCCRRRRGFRIOFQKKSEKTKETKETEESGNGNSGYNSDQKTDPTASLPKSCS 336
 DB 483 ALAFVCMRK-----IKOSCEEENAGADQDG-----EGEGSKTALQPLKHS 525
 QY 337 SDPE 340
 DB 526 KEDD 529

RESULT 12
 US-08-633-148-4
 Sequence 4, Application US/08633148
 Patent No. 5864018
 GENERAL INFORMATION:
 APPLICANT: MORSE, MICHAEL J.
 APPLICANT: NAGASHIMA, MARIKO
 APPLICANT: HOLLANDER, DORIS A.
 TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
 TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TOWNSEND & TOWNSENT & CREW LLP
 STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
 CITY: SAN FRANCISCO

STATE: CALIFORNIA
 COUNTRY: U.S.A.
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/633,148
 FILING DATE: 16-APR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY ESQ., MATTHEW B.
 REGISTRATION NUMBER: 39,787
 REFERENCE/DOCKET NUMBER: 014618-005600US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 326-2422
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 318 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-633-148-4

Query Match
 Best Local Similarity 5.8%; Score 118.5; DB 2; Length 318;
 Matches 64; Conservative 32; Mismatches 88; Indels 83; Gaps 14;

QY 15 EVIEGPONATVLSQARFNCYSSG---WKILMALSDMVVLSVRMEPIITNDRTS- 70
 DB 103 EYDSASELTA--GVNKKGTQVSBGSPAGTILSMHLDG-----KPLVNEKGVSV 151
 QY 71 ---QRTDQGNFT--SEMIHNVPEPDSGNIR---CSLQNSRLHGSAYLTVQVMELEF 120
 DB 152 KEQTRRHPTGTGLTQSELM---VTPARCGDPRPTFSCSPGDLPRHATLTAIPIQPRW 208
 QY 121 IP---SVNLVAENPP-----CEVCLPSHMTLPDLISWELGLVSHSY 162
 DB 209 EPVPLEVQLV---EPGGAAPGCTVTLTCEVPAOPS---PQIHMKMD----- 252
 QY 163 YFVPEPDLQSAVSIATLPQSNGLTCVATWKSILKARKSATVNLVIRCPDQDTGGINI 222
 DB 253 -GVPLPLPPSPVLLIPETIGPODQGYSCVATHSHGQESNAVSISITE-PGEG----- 305
 QY 223 PGVLSLPSLGFSLPTWKGVLGLAGT 249
 DB 306 -----PTAGSVGGSLGT 318

RESULT 13
 US-08-633-148-2
 Sequence 2, Application US/08633148
 Patent No. 5864018
 GENERAL INFORMATION:
 APPLICANT: MORSE, MICHAEL J.
 APPLICANT: NAGASHIMA, MARIKO
 APPLICANT: HOLLANDER, DORIS A.
 TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
 TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TOWNSEND & TOWNSENT & CREW LLP
 STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA
 COUNTRY: U.S.A.
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/633,148
 FILING DATE: 16-APR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MORRIS ESO., MATTHEW B.
 REGISTRATION NUMBER: 39,787
 REFERENCE/DOCKET NUMBER: 014618-00560005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2422
 TELEFAX: (415) 326-2400
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-633-148-2

Query Match 5.8%; Score 118.5; DB 2; Length 340;
 Best Local Similarity 24.0%; Pred. No. 0.011; Indels 83; Gaps 14;
 Matches 64; Conservative 32; Mismatches 88;

QY 15 EVIEGPNATVLKGSQARFNCYSG--WKLIIMALSDVVLVSRPMEPIITNDRTS- 70
 DB 125 EIVDSASELTA--GVPRKAVTCYSEKSYRAGTILSMHLDG-----KPLVPRNEKGVSV 173
 QY 71 ----QRYDGGNFT--SEMIHNVPSDSGNTR-----CSLQNSRLHGSAYLVQVMGELEF 120
 DB 174 KEQTRRRPELTGLTLOGLSLM---VTPARGGDPRPTSCSPGLPRHRLKTPAPLOPRVW 230
 QY 121 IP-----GVNIVVAMNEP-----CEVTCLEPSHWTLPDISWELGLVSHSSY 162
 DB 231 EPPLEEVQVLV---EPEGANVAPAGTYLTLCVPAQPS-----PQIHMKD----- 274
 QY 163 YFVPEPSDLSQAVSIALTPQSNGLTCVATWKSLSKAKSATVNLVIRCPQDTGGGINI 222
 DB 275 -GVPLPLPSPVLLIPETIGDQGTSCVATSHSHGQESRAVSISITIE-PGEG----- 327
 QY 223 PGVLSLPSLGSFSLPTWCKVGLAGT 249
 DB 328 -----PTAGSVGSGGLGT 340

RESULT 14
 US-09-651-200-20
 Sequence 20, Application US/09651200
 Patent No. 6429303

GENERAL INFORMATION:
 APPLICANT: Green et al
 TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 TITLE OF INVENTION: Polypeptides Encoded Thereby
 FILE REFERENCE: 15966-562 (CURA-62)
 CURRENT APPLICATION NUMBER: US/09/651,200
 CURRENT FILING DATE: 2000-08-30
 PRIOR APPLICATION NUMBER: 60/152383
 PRIOR FILING DATE: 1999-09-03
 PRIOR APPLICATION NUMBER: 60/172909
 PRIOR FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: 60/183578
 PRIOR FILING DATE: 2000-02-18
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 20
 LENGTH: 325
 TYPE: PRT
 ORGANISM: sus sp.
 US-09-651-200-20

Query Match 5.7%; Score 117; DB 4; Length 325;
 Best Local Similarity 22.9%; Pred. No. 0.014;
 Matches 88; Conservative 39; Mismatches 126; Indels 132; Gaps 22;

QY 3 IVIFLHSGSGNEVIEGPQNAVTKGSQARN-----C--TYSGCKL-----IMALSD 50
 DB 10 WVLISGAAS-----LK-SQAVFNEGELPCHFTNSQNSLDELVIWQDOD 55
 QY 51 MYVL--SVRPMF-PITNDRTSQ-RYDQGNFTSEMIHNVPSDSGNIRCSLQNSRLH 106
 DB 56 NLVLELYRGQEPKPNVSKTMGRSTFQA---TWTLRLHVOQKDKSGVQCFIHKGPB 112
 QY 107 GSAYL-----TVQVMGELFIPSVNLV--AENECEVTCLEPSHWTLPDISWELGLVSH 159
 DB 113 GLVPIHQMSSDLSILANSQPEINLTWHTNSVINLTC----- 151
 QY 160 SSYYFVPEPSDLSQAVSIALTPQSNGLTCVATWKSLSKAKSATVNLV---IRCPQDT 216
 DB 152 SSTQGYPEQ-----RMVMTLNTKNSSTTEHDADMKKSQNNITELYNVSIFVSLPIPET 205
 QY 217 GGINIPGVLSLPS--LGFSLP-----TWKVGGLAGTMLTPCT 257
 DB 206 --NVSTVCYLQLEPSTLFLPCNIDAKRPVQPVPHILM-----IALLVTVVW 256
 QY 258 LITRCCCCRRRCGCCGCCRCPCRRKR-----GRIOFOKK-SEKERTKET 305
 DB 257 -----CGWSPFVTLKRRKKQPPSPNECGETIKMKKASEQTKNRAEV 299
 QY 306 ETESGNEISGVN-----SDEOKTTD 325
 DB 300 HESDDAQCDVNLILKTASDNTTDD 324

RESULT 15
 US-09-540-245A-15
 Sequence 15, Application US/09540245A
 Patent No. 6270984

GENERAL INFORMATION:
 APPLICANT: Goodman, Corey
 APPLICANT: Kid, Thomas
 APPLICANT: Brose, Katja
 APPLICANT: Tessier-Lavigne, Marc
 TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 FILE REFERENCE: B98-031-3
 CURRENT APPLICATION NUMBER: US/09/540,245A
 CURRENT FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 60/065,544
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: 60/081,057
 PRIOR FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 1395
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-540-245A-15

Query Match 5.5%; Score 114; DB 4; Length 1395;
 Best Local Similarity 25.8%; Pred. No. 0.2; Indels 56; Gaps 15;
 Matches 65; Conservative 29; Mismatches 102;

QY 20 PNATVTKGSQARFNCYSG--WKLIW-----ALSDVVLVSRPMEPIITNDRTSQ 72
 DB 161 PKDTRAKKGTALLBEGPPKGIPEPTLAIKIDGVLDDLKAMSGASSRV-----R 211
 QY 73 YDQGNFTSEMIHNVPSDSGNIRCSLQNSRLHGSAYLVQVMGELEFIPSVNLVVA 129
 DB 212 IYDGN-----LLISNVEPIDEGNTKCIQNLVGTRESSYAKLIYQKYPFKPEKDDVML 267
 QY 130 ENRP-----CEVTCLEPSHWTLPDISW---ELGLVSHSSYFVPEPSDLSQAVSIALTP 182

Db 268 YGQATPFHCSVGGDDP-----PKVLWKKEGNIPVSARILHD-----EKSLEISNITP 316
QY 183 QSNGLTFQVA--TWKSLKARKSATV-----NLVIRGCPQDPTGGGINIPCVLSLPSLGFSL 236
Db 317 TDEGTVCFAHNNVGOISARASLIVHAPPFT--KRPSNKKVGLN--GVV--OLPCMASGN 371
QY 237 PT-----WGVGL 244
Db 372 PPPSVFMTKEGV 383

Search completed: April 28, 2003, 21:12:14
Job time : 18.9029 secs

XX WP1; 2002-130881/17.
 DR N-PSDB; ABR13028.
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 XX proliferative disorders, e.g. cancer and arteriosclerosis
 PS
 PS Claim 13; Fig 1; 135pp; English.
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC extracellular domains and other regulators of B7-L polypeptides comprising
 CC growth and maintenance of cancer cells based on the observation of
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{h1}.
 XX
 SQ Sequence 382 AA:

Query Match 100.0%; Score 2059; DB 23; Length 382;
 Best Local Similarity 100.0%; Pred. No. 5.5e-162;
 Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLVIFLHSGSGNVEYICGNATVNLKGSQARENCVSGQMKLIMALSDMVNLVSRPME 60
 DB 1 MGLVIFLHSGSGNVEYICGNATVNLKGSQARENCVSGQMKLIMALSDMVNLVSRPME 60
 QY 61 PIITNDRTSGRYDGGNFTSEMIHNVPSDSGNTRCSLQNSRLGSAVLYVWGELEF 120
 DB 61 PIITNDRTSGRYDGGNFTSEMIHNVPSDSGNTRCSLQNSRLGSAVLYVWGELEF 120
 QY 121 IPSVNLVAEKEPEVCYCLPSHWTLPDISWEIGLLVSHSYFYFEPEDLDQASVIAL 180
 DB 121 IPSVNLVAEKEPEVCYCLPSHWTLPDISWEIGLLVSHSYFYFEPEDLDQASVIAL 180
 QY 181 TPQSNGLTFCVATWMSLKARKSATVNLVYIRCPDQDGGGINIPVGLSSIPSLGSLPTWG 240
 DB 181 TPQSNGLTFCVATWMSLKARKSATVNLVYIRCPDQDGGGINIPVGLSSIPSLGSLPTWG 240
 QY 241 KVGGLAGTMTLLPTCTLTIRCCCRRCGCCGCCRCFCRCRRGRGRIOPKKSEK 300
 DB 241 KVGGLAGTMTLLPTCTLTIRCCCRRCGCCGCCRCFCRCRRGRGRIOPKKSEK 300

QY 301 TNKETESGNSGNSDQKTTDTASLPKSCSSPEQNRSSCGPPHROAROPRRP 360
 DB 301 TNKETESGNSGNSGNSDQKTTDTASLPKSCSSPEQNRSSCGPPHROAROPRRP 360
 QY 361 ASHPQASFNLAPEKSNNTTV 382
 DB 361 ASHPQASFNLAPEKSNNTTV 382
 PS
 PS Claim 13; Fig 3; 135pp; English.
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC extracellular domains and other regulators of B7-L polypeptides comprising
 CC growth and maintenance of cancer cells based on the observation of
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Human B7-like protein, B7-L_{h3}.

Query Match	Best Local Similarity	73.5%	Score 1514	DB 23	Length 377
Matches 284	Conservative 12	Mismatches 20	Indels	6	Gaps 2
QY	9	GSSGSGNEVEYEGPQNAATVLKGSQARFNCTVSGMKLIMMALSDMWVLSVRMEPIITNDRF	68		
DB	13	GSSGSGNEVEYEGPQNAATVLKGSQARFNCTVSGMKLIMMALSDMWVLSVRMEPIITNDRF	68		
QY	69	TSQRYDQGGNFESEMIINHPVSDSGNIRCSLQNSRLHGSAYLVTVGMELFIPSYNLVY	72		
DB	73	TSQRYDQGGNFESEMIINHPVSDSGNIRCSLQNSRLHGSAYLVTVGMELFIPSYNLVY	128		
QY	129	AENECEVETCLPSHWTRLPLDIEMELGLVSHSSYFVPEPSDLOSASTLATPQSNCTL	188		
DB	133	AENECEVETCLPSHWTRLPLDIEMELGLVSHSSYFVPEPSDLOSASTLATPQSNCTL	192		
QY	189	TCVATWKSLSARSKASATVNLVIRCPDQTGGGINIPGVYLSLPSLGFSLPTWKGYSGLAG	248		
DB	193	TCVATWKSLSARSKASATVNLVIRCPDQTGGGINIPGVYLSLPSLGFSLPTWKGYSGLAG	252		
QY	249	TMLTLPCTCTLTIRCCCRRCGCCGCCCCGCCCRKRGFIPOKSKSEKRTNKEETE	308		
DB	253	TMLTLPCTCTLTIRCCCRRCGCCGCCCCGCCCRKRG---NLKRRRQTKLRLQKVEK	309		
QY	309	SGNENSGYNSDEKTKTDVNASLP	330		
DB	310	T---PATIQMKRKPQTPLSLP	328		
RESULT 5					
ID	ABG28169	standard; Protein; 463 AA.			
AC	ABG28169;				
DT	18-FEB-2002	(first entry)			
DE	Novel human diagnostic protein #28160.				
KV	Human: chromosome mapping; gene mapping; gene therapy; forensic;				
XX	food supplement; medical imaging; diagnostic; genetic disorder.				
OS	Homo sapiens.				

CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m1}.

SO Sequence 370 AA:

Query Match 44.9%, Score 924.5; DB 23; Length 370;
 Best Local Similarity 50.8%, Pred. No. 3,6e-66;
 Matches 197; Conservative 50; Mismatches 104; Indels 37; Gaps 6;

QY 3 LVIF--LHGSGSGNEVTEGPNATVVLKSGQAFNCTVSGQMKLIMVALSDMVLSVRPME 60
 DB 12 LVIIAQLTASSSSVQIIEGPNATVVLKSDSAHFNCTVTHGKMLMTLQMVLSLTGQ 71
 QY 61 PIIINDRTSORYDQGNFTSEMIINVEPSDGNIRCSLQNSRLGSAVLVQVNGELF 120
 DB 72 PIIINRRTYASVNSTDSFTSELIHDVQSDSGVQCSLQNSHGFSAFLSVQVNGELN 131
 QY 121 IPSVNLVAENPECEVTCLEPSHMTRLPDISWELGLVSHSYFVEPEPSDLQSAVSIIL 180
 DB 132 IPSNNLIVTEGECNVTCAVGMTSLPDISWELEVPVSHSYNSFLEGEENRVLSTLDL 191
 QY 181 TPDSNGITLVATWKLARKRSATVNLVIRCPDPTGGGGINIPVLSLPSLGSFLPTWG 240
 DB 192 TPLNGETLCVAMELKDQASKSLVNLVYQPPD-----SIGEGPALPTWA 239
 QY 241 KVLGLAGTMLTLPCTLIRCCCRKCCGCCGCCRCFCRRKGRFRIDFQ---KKSE 297
 DB 240 IILAAVAFSLILILVILITFCCC-----CASRREKEEPTVQEIIRKSA 284
 QY 298 KKKTKK---ETETSGNENSGYNSDCKTDTLASLPPKSCSSPDEQRNSCGEPHQRAD 354
 DB 285 NKRTKADPEKRLKSGKENYGYSSDEAKAQTASLPKSAVSLPEKRSSL--PYOEIN 342
 QY 335 QRRPPASHPOASFNLASPEKVSNTTVY 382
 DB 343 KHQGPATHPRVSFDIASPOKVRNVTLV 370

RESULT 7

AAU75547
 XX ANU75547 standard; Protein; 631 AA.

AC AAU75547;

DT 23-APR-2002 (first entry)

DE Rat B7-like protein, B7-1.

KM anti-tumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KM anti-inflammation; dermatologic; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antidiabetic; antiallergic;
 KM antiautismic; nephrotoxic; antibacterial; virocidic; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; nephropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder.

OS Rattus rattus.

PN W0200200710-A2.

XX 03-JAN-2002.

PD 28-JUN-2001; 2001MO-US20719.

PR 28-JUN-2000; 2000US-214512P.

PR 28-NOV-2000; 2000US-0729264.

PA (AMGE-) AMGEN INC.

PI Welcher AA, Samlento UM, Schultz HJ, Chute HT;
 DR WPI; 2002-130881/17.

XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 PS Disclosure; Fig 8; 135pp; English.

CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular stenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or all sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC pemphigoid (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC thrombocytopenias), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of rat B7-1.

SO Sequence 631 AA;

Query Match 28.3%, Score 583.5; DB 23; Length 631;
 Best Local Similarity 26.9%, Pred. No. 1.2e-19;
 Matches 173; Conservative 43; Mismatches 111; Indels 317; Gaps 11;

QY 3 LVIF--LHGSGSGNEVTEGPNATVVLKSGQAFNCTVSGQMKLIMVALSDMVLSVRPME 21
 DB 12 LVIIAQLTASSSSVQIIEGPNATVVLKSDSAHFNCTVTHGKMLMTLQMVLSLTGQ 71
 QY 22 -----NATVLKSGQAFNCTVSGQMKLIMVALSDMVLSVRPMEPIITNDRF---- 68
 DB 72 VGVQSKSVRNVTYVKDSEAHFNCTVTHGKMLMTLQMVLSLTGQPIITNDRF 131
 QY 69 -----TSORYDQGNFTSEMI 84
 DB 132 LSCDYKRCSEFSIHRLTWCKHDKVLSVSGVPEWPKKRTTYASVNSTDSFTSELI 191
 QY 85 IHNVEPSDGNIRCSLQNSRLGSAVL----- 111
 DB 192 IHDVQSDSGVQCSLQNSHGFSAFLSVQVNDIANNVSESLGLILSDRGTYCVQRY 251
 QY 112 -----TVQVNGELFISVNLVAENPECEVTCLEPSHMTRLPDISWELGLVSHS 160
 DB 252 EGSYVYKHLITVEVWGLTNLIPSNLIVTEGECNVTCAVGMTSLPDISWELEVPVSHS 311

QY 161 -----SYFVPEPSDQ 172
 Db 312 LSVRADEPTPTNITEYGNPSADIRITCFASGCGPKPRRLSNLENGRELNSYNGFLEPGNFM 371
 QY 173 SAVSIIALTPQSNGLTLCVATFKSLKARKSATVNL----- 207
 Db 372 RVLSVLDITFGNGLTLCVAFELKDLQASKSLVNLGINTTISQDPESELYTISQDLFNA 431
 QY 208 -----TYRCQDPTGGGGINIPGVLSLPSLPGSFLPTMGKGLGL 246
 Db 432 TYDHTDCEFIYGDPAHVQSNTVQPPD-----SIEGKSPALPTAIIILLAV 479
 QY 247 AGTML-----TPT 255
 Db 480 AFSLLIIIIYIIITFWKRPEDPEKQVPEFAMGPDVAKAIIFIAITVIAVIA 539
 QY 256 CTLTRCCGCCRRRCGCCGCCRCRRKGFRIQV-KKSEKKNK---ETETESGN 311
 Db 540 IAIIFCCCSASR-----REKEESTYQNEIRKSANMRKNKADPETKLSGK 585
 QY 312 EMSGYNDEQKTDYASLPPKSCSSDEQORNSCGPPHORADQ 355
 Db 586 ENYGSYSECTIVKRRRCRRRNRNSRETNNIYIGPEAAEQ 629

RESULT 8
 AAU75545
 ID AAU75545 standard; Protein: 270 AA.
 AC AAU75545;
 XX
 DT 23-APR-2002 (first entry)
 DE Mouse B7-1-like protein, B7-L-m2.
 XX
 KW Mouse; B7-1-like protein; B7-L; antiinfectility; gynaecological;
 KW antitumor; cytostatic; immunosuppressive; antihistritic; antineumatic;
 KW antiinflammatory; dermatological; antiporiatic; neuroprotective;
 KW antidiabetic; haemostatic; antihydroly; anticancer; cancer;
 KW antidiabetic; nephrotoxic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 XX
 OS Mus musculus.
 XX
 PN MO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Weicher AA, Sarmiento UM, Schultz HU, Chute HT;
 DR WPI; 2002-130881/17.
 DR N-PSDB; ABK13033.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 13; Fig 6; 135pp; English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and

CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or alloimmunisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L-m2.
 XX
 SQ Sequence 270 AA;
 QY Query Match 28.0%; Score 575.5; DB 23; Length 270;
 Best Local Similarity 35.1%; Pred. No. 1.8e-39;
 Matches 136; Conservative 42; Mismatches 73; Indels 137; Gaps 6;

QY 3 LVTF--LHSGSGENVEYEGQNAITYKGSQARFNCVSGKLMALSDMYVLSVREME 60
 Db 12 LVTLAQLTASGSSVQIIEBPQNTVYKQSEAFNCTVTHGKLLMWTINQVYVLSLTQQG 71
 QY 61 PITTNDFTSORDQGNFTSEMIHNVERSDSGNTRCSLQNSRLGSAVLYTVOMGLF 120
 Db 72 PIIITNRFYTAISNDSFISLIIHDVQPSDGSVQCSLQNSHGFSAFLSVQ----- 125
 QY 121 IPSVNLVVAENDECEVTCIPSHWTRLPDISWELGLVSHSYFYFPEPSDQSAVSLAL 180
 Db 126 ----- 125
 QY 181 TPQSNGLTLCVATFKSLKARKSATVNLTVIRCPQDTGGGINIPGVLSLPSLPGSLPTWG 240
 Db 126 -----DSIEGKSPALPTWA 139
 QY 241 KVGGLAGTMLPPTCTLTIRCCGCCGCCRCRCKRRKGFRIQV-----KKSE 297
 Db 140 IILAVAFSLIIIIYIIIFCC-----CASRREKEESTYQNEIRKSA 184
 QY 298 KEKTNK--ETETESGNENSGYNSDDEKTTDTASLPPKSCSSDPPQORNSCGPPHORAD 354
 Db 185 NMRTNKAADPETKLSGKENGYSDEKAAQITSLPKSAEYSLPKRRSSSL--PYQELN 242
 QY 355 QRPPEASHPOASFNLASPEKVSNTTVV 382
 Db 243 KHQEPALHPVSPFDIASPQKVRNVTIV 270

RESULT 9
 AAU75546
 ID AAU75546 standard; Protein: 223 AA.
 AC AAU75546;
 XX

DT 23-APR-2002 (first entry)
 XX
 DE Mouse B7-like protein, B7-L_{m3}.
 XX
 XX Mouse: B7-like protein; B7-L; antiinfertility; gynaecological;
 KM antitumour; cytostatic; immunosuppressive; antiarthritic; antineumatic;
 KM antiinflammatory; dermatological; antiporiatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
 KM antistimatic; nephrotropic; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; nephropathy; skin disorder;
 XX endocrinopathy; lymphoproliferative disorder.
 OS Mus musculus.
 XX
 PN WO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PD 28-JUN-2001; 2001WO-US20719.
 XX
 PF 28-JUN-2000; 2000US-214512P.
 XX
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 DR WPI: 2002-130881/17.
 DR N-PSDB: ABK1034.
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 XX Claim 13; Fig 7; 135pp. English.
 PS
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC extracellular disorders. Antibodies, soluble proteins comprising
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of hematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m3}.

XX
 SQ Sequence 223 AA:
 Query Match 24.0%; Score 494; DB 23; Length 223;
 Best Local Similarity 31.9%; Pred. No. 7,7e-33;
 Matches 123; Conservative 37; Mismatches 47; Indels 178; Gaps 7;
 QY 3 LVYF--LHSGSGGNEVIEGPONATVYLGSGA RPNCTVSGGKLLMALSDMYVLSVRME 60
 DB 12 LVYLAQLTASSGSSVYIIEGPONATVYLGSGA RPNCTVSGGKLLMALSDMYVLSVRME 60
 QY 61 PIIITDRFTSCRYTDOGGNFTSEMIINHNVEPSDSGNIRCSLQNSRLHGSAYLTVOVGELP 120
 DB 72 PIIITNRRFTVASYNSTDSFIISLLIHVDYQPSDSSGVCSLQNSHGRSAYLSVQ----- 125
 QY 121 IPSVNLVVAENPECVTCLPSHWTRLPDISMELGLVSHSSYTFVPEPSDLSGSAVSITLAL 180
 DB 126 -----ESTY-----QNEI----- 133
 QY 181 TPQSNGLTVCAVWMSLKARKSATVNLTVIRCPQDTGGGINIPGLVSLPSLGFSLPWTG 240
 DB 134 -----RKSANM----- 139
 QY 241 KVGGLAGTMLTPTGTLTRCCCRRCGCCGCCCCRCCECCRRKRGRPIQKRSSEK 300
 DB 140 -----R 140
 QY 301 TNK---ETETESGNGNSGNSYNSDEQKTTDTASLPKSCSSSDPEORNSGCPPHORADQRP 357
 DB 141 TNKADPEFKLKSNGENYSGSSDEKAAQTASLPKSAEYSLPEKRSLSL--PYQELNRHQ 198
 QY 358 PRPASHPOASPNLASPEKYSNTTVV 382
 DB 199 PGPAIHPRVSFIDASPOKVRNVTLV 223
 RESULT 10
 AAB81925
 ID AAB81925 standard; protein: 404 AA.
 XX
 AC AAB81925;
 XX
 DT 15-JUN-2001 (first entry)
 XX
 DE Extracorporeal circulation material receptor protein.
 XX
 KM Extracorporeal circulation; carbonyl stress product; receptor;
 KM diabetes; vascular lesion; excretory dysfunction.
 XX
 OS unidentified.
 XX
 PN WO200118060-A1.
 XX
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000WO-JP06172.
 XX
 PR 08-SEP-1999; 99JP-0254463.
 XX
 PA (TORA) TORAY IND INC.
 XX
 PI Shimizu S, Kubota M, Akiyama H, Usui M;
 DR WPI: 2001-290314/30.
 XX
 PT Material for extracorporeal circulation, applicable in selective
 PT elimination of diabetic complication factors such as carbonyl stress
 PT products caused by abnormally promoted carbonyl stress from excretory
 PT dysfunction in vascular lesions
 XX
 PS Claim 1; Page 31-32; 36pp; Japanese.
 XX
 CC The present invention describes a material for extracorporeal circulation

which is made from a water-insoluble carrier immobilized with a protein having the sequence shown here. The materials of the invention, including adsorbents, are for extracorporeal circulation, which are applicable in the selective elimination of diabetic complication factors from a body fluid, and are therefore useful in treating vascular lesions like arteriosclerosis due to carbonyl stress products caused by abnormally promoted carbonyl stress from excretory dysfunction.

Sequence 404 AA:

Query Match 7.2%; Score 147.5; DB 22; Length 404;
Best Local Similarity 22.6%; Pred. No. 0.00076;
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

15 EYIEGPNATVILKGSQARNCVTSGQ---WKLIMALSDMYLVSVPMEPIITNDRTFS- 70
125 EIVDSASELTA--GVPNKVGTCSGSPAGTILSMHLDG-----KPLVPMKGVSV 173
71 ---QRYDGGNFT--SEMIHNVPSDSGNIR---CSLONSRLHGSAYLTVQVMGELEF 120
174 KEOTRRHPETGLFTLOSSEM--VTPARGGDRPPTSCSFSPGLPRHRALFTAPIORRW 230
121 IP---SVNLVVAENP-----CEVTLCPSHWTRLPDISMELGLVSHSSY 162
231 EPPVLEEVQLV---EPREGAVAPGGVTLLICEVPAQPS---PQIHMKD----- 274
163 YFVPEPSDIQSASVILATPQNSGTLTCVATWKSILAKRSATVNLTVRCPODTGGGINI 222
275 -GYPLPLPSPVILPEIPGODGYSCVATHSSHGQPSRAVSISITIE-PEEBG----- 327
223 PGVLSSLPISLGFSLPTWKGVLGLAGTMTLT-----PTCTLTTRCCCRRCGCCNCC 275
328 -----PAGSVGGSGGLTALALGILGGLTALLIGVILMQR----- 366
276 CRCCFCRRKRGFRIOFQKSEKERT--NKETETSGNNSGNSGNDKOTTTTASLPKPS 333
367 -----QRRGEERKAPENOEERAEIN----- 389
334 CESSDPEQRNSSCGPP 349
390 -QSEEPAGESSTGCP 404

RESULT 11

AAE23219
ID AAE23219 standard; Protein; 404 AA.

AAE23219;

27-NOV-2002 (first entry)

Human receptor for advanced glycosylation end product (RAGE) protein.

Human: Receptor for advanced glycosylation end product; RAGE; cardiact; tissue growth; neointimal formation; blood vessel; restenosis; diabetes; myocardial infarction; angioplasty; peripheral vascular surgery; angina; transgenic animal; acute thrombotic stroke; venous thrombosis.

Homo sapiens.

WO200230889-A2.

18-APR-2002.

12-OCT-2001; 2001WO-US32036.

13-OCT-2000; 2000US-0687528.

(UYCO) UNIV COLUMBIA NEW YORK.

Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;

WPI; 2002-426260/45.

N-PSDB; AAD36952.

Inhibiting new tissue growth or neointimal formation in blood vessels of subject suffering from diabetes, stroke and preventing restenosis, comprises administering inhibitor of receptor for advance glycation end product -
Disclosure; Page 16; 43pp; English.

The invention relates to a method for inhibiting new tissue growth or neointimal formation in blood vessels in a subject that has experienced blood vessel injury and preventing exaggerated restenosis in a diabetic subject. The method comprises administering an inhibitor of receptor for advanced glycation/glycosylation end product (RAGE), so as to inhibit new tissue growth or neointimal formation in subject's blood vessels and preventing restenosis in the subject. The method is useful for inhibiting new tissue growth or neointimal formation in blood vessels in a subject like non-human animal, a transgenic non-human animal or a human suffering from diabetes, acute thrombotic stroke, venous following angioplasty or angina, myocardial infarction, abrupt closure following peripheral vascular surgery stent placement, or thrombosis as a result of peripheral vascular surgery. The method is also useful for preventing restenosis and for determining whether a compound inhibits new tissue growth in a blood vessel in a subject. The present sequence is human receptor for advanced glycosylation end product (RAGE) protein.

Sequence 404 AA:

Query Match 7.2%; Score 147.5; DB 23; Length 404;
Best Local Similarity 22.6%; Pred. No. 0.00076;
Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

15 EYIEGPNATVILKGSQARNCVTSGQ---WKLIMALSDMYLVSVPMEPIITNDRTFS- 70
125 EIVDSASELTA--GVPNKVGTCSGSPAGTILSMHLDG-----KPLVPMKGVSV 173
71 ---QRYDGGNFT--SEMIHNVPSDSGNIR---CSLONSRLHGSAYLTVQVMGELEF 120
174 KEOTRRHPETGLFTLOSSEM--VTPARGGDRPPTSCSFSPGLPRHRALFTAPIORRW 230
121 IP---SVNLVVAENP-----CEVTLCPSHWTRLPDISMELGLVSHSSY 162
231 EPPVLEEVQLV---EPREGAVAPGGVTLLICEVPAQPS---PQIHMKD----- 274
163 YFVPEPSDIQSASVILATPQNSGTLTCVATWKSILAKRSATVNLTVRCPODTGGGINI 222
275 -GYPLPLPSPVILPEIPGODGYSCVATHSSHGQPSRAVSISITIE-PEEBG----- 327
223 PGVLSSLPISLGFSLPTWKGVLGLAGTMTLT-----PTCTLTTRCCCRRCGCCNCC 275
328 -----PAGSVGGSGGLTALALGILGGLTALLIGVILMQR----- 366
276 CRCCFCRRKRGFRIOFQKSEKERT--NKETETSGNNSGNSGNDKOTTTTASLPKPS 333
367 -----QRRGEERKAPENOEERAEIN----- 389
334 CESSDPEQRNSSCGPP 349
390 -QSEEPAGESSTGCP 404

RESULT 12

AAU77543
ID AAU77543 standard; Protein; 404 AA.

AAU77543;

05-JUN-2002 (first entry)

Human receptor for advanced glycosylation end product (RAGE).

Receptor for advanced glycation end product; RAGE; receptor; amyloid beta peptide; blood-brain barrier; neurovascular stress;

KW cerebral vasoconstriction suppressor; cerebral blood flow enhancer;
 KW cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;
 KW Alzheimer's disease; Down's syndrome; head trauma; stroke; human.

OS Homo sapiens.

PN W0200214519-A1.

PD 21-FEB-2002.

PF 14-AUG-2001; 2001WO-US25416.

PR 14-AUG-2000; 2000US-0638648.

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Stern DM, Schmidt AM, Yan SD, Zlokovic B.

DR WPI; 2002-257610/30.

DR N-PSDB; ABK10856.

PT Ameliorating neurovascular stress and decreasing cerebral
 PT vasoconstriction in subject suffering from chronic/acute cerebral
 PT amyloid angiopathy, by administering inhibitor of receptor for advanced
 PT glycation endproduct

PS Disclosure; Page 16; 68pp; English.

XX The invention describes a method of ameliorating neurovascular stress,
 CC and decreasing cerebral vasoconstriction in subject suffering from
 CC chronic or acute cerebral amyloid angiopathy, comprising administering
 CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).
 CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain
 CC barrier, thus decreasing cerebral vasoconstriction and increasing
 CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
 CC subject, decreasing cerebral vasoconstriction in a transgenic non-human
 CC animal (preferably, transgenic mouse overexpressing mutant human amyloid
 CC beta precursor protein) or a human, suffering from chronic or acute
 CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
 CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy
 CC in a subject, where the neurovascular stress is caused by Alzheimer's
 CC disease, aging, Down's syndrome, head trauma or stroke. This is the
 CC amino acid sequence of human receptor for advanced glycation end
 CC product (RAGE) described in the invention.

XX Sequence 404 AA;

Query Match 7.2%; Score 147.5; DB 23; Length 404;

Best Local Similarity 22.6%; Pred. No. 0.00076;

Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

QY 15 EVIEGPNATVLEKGSQARFNCVTSOG---WKLIMMALSDMWLVSVRPMPEIITNDRTS- 70

DB 125 EIVDSASELTA--GVNKNKVCVSEGSYPAGTILSHLDG-----KPLVPEKGVSV 173

QY 71 ----ORYDGGNFT--SEMIINHVPSDSGNIR---CSLONSRLHGSAYITVQVMGELF 120

DB 174 KEQTRRHPEFTGLTQSELM---VTPARGDPRPTSCFSFGPLRRHRLRTAPIQPRVW 230

QY 121 IP---SVNLVAENEP-----CEVTCLPSHMTLRPLDISWELGLLVSHSY 162

DB 231 EPPVLEEVQLV---EEGGAVALPGGTVILTCVPAOPS-----PQIHMKD----- 274

QY 163 YFVPEBDSLOANSIILALPQSNGLTCVATWKSILKARSAVNLTVIRCPDGTGGINI 222

DB 275 -GVPLPPLPSPVILPEIGPODGTYSVATHSHSGPSRAVSIISIE-PEEB----- 327

QY 223 PGVLSLPSLGSFLPTMGKVGGLACTMILT-----PTCLTLTRCCCCRRRCGCCNCC 275

DB 328 -----PTAGSYGSGGLTALALGILGGLTALLIGVILMQR----- 366

QY 276 CRCCFCRRKRGRIQFORKSEKRT--NKETETESGNENSGYNSDEKTTDTASLPPKS 333

DB 367 -----QRGEERKAPENQEEBERALN----- 389

QY 334 CESDPEQRNSGCPP 349

DB 390 -QSEEPGAGESTGCP 404

RESULT 13

AAM48745

ID AAM48745 standard; protein; 404 AA.

AC AAM48745;

DT 02-APR-2002 (first entry)

XX Human RAGE protein SEQ ID NO 1.

DE Human: RAGE; receptor for advanced glycosylated endproduct; receptor;

KW antidiabetic; neuroprotective; cyostatic; antiinflammatory; vasotropic;

KW nephrotropic; dermatological; antiarteriosclerotic; nootropic; diabetes;

KW Alzheimer's disease; cancer; inflammation; kidney failure;

XX systemic lupus; nephritis; erectile dysfunction; atherosclerosis.

OS Homo sapiens.

PN W0200192892-A2.

PD 06-DEC-2001.

PF 30-MAY-2001; 2001WO-US17447.

PR 30-MAY-2000; 2000US-207342P.

PR 05-MAR-2001; 2001US-0799152.

PA (TRAN-) TRANS TECH PHARMA.

XX Shahbaz M;

PI WPI; 2002-114372/15.

DR Detecting a receptor for advanced glycosylated endproducts (RAGE)

XX modulators, for treating e.g., cancer, diabetes or inflammation,

PT comprises measuring the amount of bound anti-RAGE antibody

XX Claim 1; Fig 2; 49pp; English.

PS The invention relates to detecting receptor for advanced glycosylated

XX endproducts (RAGE) modulators comprises determining the amount of RAGE

CC protein or its fragment bound to the pre-adsorbed ligand by measuring the

CC amount of anti-RAGE antibody bound to the solid surface. The method is

CC useful for rapid, high-throughput identification of compounds that

CC modulate RAGE. The compounds are useful for treating symptoms of diabetes

CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's

CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis

CC or inflammatory lupus nephritis, erectile dysfunction and

XX atherosclerosis.

XX Sequence 404 AA;

Query Match 7.2%; Score 147.5; DB 23; Length 404;

Best Local Similarity 22.6%; Pred. No. 0.00076;

Matches 85; Conservative 40; Mismatches 114; Indels 137; Gaps 18;

QY 15 EVIEGPNATVLEKGSQARFNCVTSOG---WKLIMMALSDMWLVSVRPMPEIITNDRTS- 70

DB 125 EIVDSASELTA--GVNKNKVCVSEGSYPAGTILSHLDG-----KPLVPEKGVSV 173

QY 71 ----ORYDGGNFT--SEMIINHVPSDSGNIR---CSLONSRLHGSAYITVQVMGELF 120

DB 174 KEQTRRHPEFTGLTQSELM---VTPARGDPRPTSCFSFGPLRRHRLRTAPIQPRVW 230

QY 121 IP---SVNLVAENEP-----CEVTCLPSHMTLRPLDISWELGLLVSHSY 162

33 (1977) 1 111

senile dementia; migraine; neurasthenia; neuropathy; autoimmune disease; Alzheimer's disease; kw

diagnosis; gene therapy; pancreatitis; inflammation; tumour; cancer

diagnosis; gene therapy

KW neural degeneration; antiinflammatory; cytostatic; nootropic;
 KW immunosuppressive; antiparkinsonian; neuroprotective; antimigraine;
 KW anticonvulsant.
 OS Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= Signal_peptide
 FT 19..594
 FT /label= Mature_protein
 XX
 PN MO200198360-A2.
 XX
 PD 27-DEC-2001.
 XX
 XX 22-JUN-2001; 2001WO-US19904.
 XX
 XX 22-JUN-2000; 2000US-213611P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (CARU/) CARULLI J P.
 PA (LUKA/) LUKASHIN A V.
 PA (KILB/) KILBURN D R.
 PA (SUNC/) SUN C.
 XX
 PI Carulli JP, Lukashin AV, Kilburn DR, Sun C;
 XX
 DR MPI: 2002-329171/36.
 DR N-PSDB; ABL3877.
 XX
 PT Novel nucleic acid sequence encoding a member of immunoglobulin
 PT superfamily, designated GP354, useful for the treatment of Alzheimer's
 PT disease, Parkinson's disease, senile dementia, migraine and epilepsy -
 XX
 PS Claim 17; Fig 8; 163pp; English.
 XX
 CC The present sequence is the protein sequence of human pancreatic
 CC GP354, a novel member of the immunoglobulin superfamily. GP354 is
 CC a pancreas-enriched integral membrane protein, which is also
 CC detected at low levels in central nervous system (CNS) tissue. The
 CC protein structure and tissue distribution indicate a role for GP354
 CC in cell-cell recognition, binding, signalling and adhesion events
 CC in the pancreas and CNS. The invention provides GP354 polypeptides
 CC antibodies and related diagnostic and therapeutic methods. Claimed
 CC compositions comprising a GP354 nucleic acid or GP354 polypeptide
 CC are used in the treatment of pancreatic injury and abnormal or
 CC disease conditions that relate to the pancreas, such as acute or
 CC chronic pancreatitis, pancreatic inflammation, pancreatic necrosis,
 CC exocrine insufficiency, pancreatic endocrine and hormonal imbalance,
 CC pancreatic tumours and associated cancers, and autoimmune disorders,
 CC which affect the pancreas. They are also used in the treatment of
 CC an injury to the CNS and abnormal or disease conditions that relate
 CC to the CNS, including Alzheimer's disease, Parkinson's disease,
 CC senile dementia, migraine, epilepsy, neuritis, neurasthenia,
 CC neuropathy, neural degeneration and neural tumours (all claimed).
 CC
 XX
 SQ Sequence 594 AA;
 Query Match 6.6%; Score 136; DB 23; Length 594;
 Best Local Similarity 23.8%; Pred. NO. 0.011;
 Matches 55; Conservative 40; Mismatches 106; Indels 30; Gaps 10;
 QY 3 LVIFLHSGSGGNEVIEGPNATVILKSGQARFNCYSGKLMWALSDMVLVSVPMPEI 62
 DB 10 LFCFLRRAGPSPHPLQOEDEVLVIGEBARLPCALGATWGLVQWTKSGIALGGR----- 64
 QY 63 ITNDRFTSORVYDGGNFMPS---EMTIHVPSDSGNIRCSLQNSRLHG-SAYLTV----- 113
 DB 65 ---DLPGMSRYWISGNANAGOHDLHTRPELEDEASYEQATGAGLRSPALHVLVPPPE 121
 QY 114 --QVNGELFIPBSVNLVAENPECEVTCLPSHWTR-LPDISM-ELGLLVSHSY--FVPE 167

DB 122 APQVLGG---PSYSIVA--GVFANLTCNSRGDARPTPELWFRDGVLLDGATFHQTLIKE 176
 QY 168 --PSDLQSAVSIATLTPQSNSTLTVCVATWKSILKARKSATVNLVIRCPQDT 216
 DB 177 GTPGSVESTLTLTLPESHDDGATFVCARNSQALPTGRDAITLISIQYPEVT 227

Search completed: April 28, 2003, 18:14:28
 Job time : 43.0676 secs

Query Match
 Best Local Similarity 81.0%; Score 1692; DB 4; Length 315;
 Matches 314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14 SSGNEVIEGPNQARVYKLSQARFNCYVSGQKLMALSDMVYLSYRPMETITNDRET 73
 DB 1 SSGNEVIEGPNQARVYKLSQARFNCYVSGQKLMALSDMVYLSYRPMETITNDRET 60

QY 74 SQRVDGNGFTSEMIHNEVPSDNGINCSLONSRLHGSATLVQVWGLFIPSVNLVA 133
 DB 61 SQRVDGNGFTSEMIHNEVPSDNGINCSLONSRLHGSATLVQVWGLFIPSVNLVA 120

QY 134 ENEPEVYCLPSHMTWLPDISWELGLVSHSSYYFVPEPSDLSQAVSILALTPQSNCTLT 193
 DB 121 ENEPEVYCLPSHMTWLPDISWELGLVSHSSYYFVPEPSDLSQAVSILALTPQSNCTLT 180

QY 194 CVATWKSILARKSATVNLVIRCPDPTGGGINTIPGVLSPSLGSLPTWKGVLGLAGT 253
 DB 181 CVATWKSILARKSATVNLVIRCPDPTGGGINTIPGVLSPSLGSLPTWKGVLGLAGT 240

QY 254 MLTPPTCTLTIRCCCRRCGCCGCCRCRRCRRGRFIOFQKSEKTKETKETEES 313
 DB 241 MLTPPTCTLTIRCCCRRCGCCGCCRCRRCRRGRFIOFQKSEKTKETKETEES 300

QY 314 GNENSGYNSDQKTT 328
 DB 301 GNENSGYNSDQKTT 315

RESULT 2
 Q9DBG2 PRELIMINARY; PRT: 270 AA.

AC Q9DBG2;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 2010003D20R1K protein (RIKEN cDNA 2010003D20 gene).
 GN 2010003D20R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker Y., Wilmink L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK008060; BAR25436.1;
 DR EMBL: BC004806; AA04806.1;
 DR MGD: MGI:1919306; 2010003D20R1K.

DR InterPro: IPR003599; IG.
 DR InterPro: IPR003006; IG_MHC.
 DR SMART: SM00409; IG; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN; 1.
 SQ SEQUENCE 270 AA; 29604 MW; A39C273DA950DFE CRC64;

Query Match
 Best Local Similarity 34.6%; Score 566.5; DB 11; Length 270;
 Matches 131; Conservative 42; Mismatches 71; Indels 135; Gaps 5;

QY 14 SSGNEVIEGPNQARVYKLSQARFNCYVSGQKLMALSDMVYLSYRPMETITNDRET 73
 DB 21 SSGNEVIEGPNQARVYKLSQARFNCYVSGQKLMALSDMVYLSYRPMETITNDRET 80

QY 74 SQRVDGNGFTSEMIHNEVPSDNGINCSLONSRLHGSATLVQVWGLFIPSVNLVA 133
 DB 81 SQRVDGNGFTSEMIHNEVPSDNGINCSLONSRLHGSATLVQVWGLFIPSVNLVA 125

QY 134 ENEPEVYCLPSHMTWLPDISWELGLVSHSSYYFVPEPSDLSQAVSILALTPQSNCTLT 193
 DB 126 ENEPEVYCLPSHMTWLPDISWELGLVSHSSYYFVPEPSDLSQAVSILALTPQSNCTLT 125

QY 194 CVATWKSILARKSATVNLVIRCPDPTGGGINTIPGVLSPSLGSLPTWKGVLGLAGT 253
 DB 126 CVATWKSILARKSATVNLVIRCPDPTGGGINTIPGVLSPSLGSLPTWKGVLGLAGT 148

QY 254 MLTPPTCTLTIRCCCRRCGCCGCCRCRRCRRGRFIOFQKSEKTKETKETEES 307
 DB 149 MLTPPTCTLTIRCCCRRCGCCGCCRCRRCRRGRFIOFQKSEKTKETKETEES 193

QY 308 ETEESGNGYNSDQKTTDPTASLPKSCSSDPEQNSCGPPHODRPPRASH 367
 DB 194 ETKKSKSEKVEYGSDEPAKQATASLPKSAEVSLEPKRSSL--PYDELKHKQGPAT 251

QY 368 PQASFNLASPEKVSNTTVV 386
 DB 252 PQASFNLASPEKVSNTTVV 270

RESULT 3
 Q9DBG2 PRELIMINARY; PRT: 168 AA.

AC Q9DBG2;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 4931420D14R1K protein.
 GN 4931420D14R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker Y., Wilmink L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

QY	260	CTITFRCCOORRRCCCNC--CGRCCGCCRRRGRIOPKRSKEKETNKETESGENS	318
Dd	55	CSLRSCCCCCRCRC-CYCRCRC--CCSRRRRFRRTLRVBDPRFQ--ITEKGDSL	107
QY	319	GYNDEKRTD-----TASLPKCSDESDEQRNSSCGPHQARQRPRAASHPQ	367
Dd	110	QRTRRQLTSOLELIEPEPTWALEPSETVAFFSHKANVSDF-----EVPPLCDSDPF	165
QY	370	ASFNLAS	376
Dd	166	PNGDLAS	172

DT	01-NOV-1996 (TREMBLrel. 01)	Created)
DT	01-JAN-1998 (TREMBLrel. 05)	Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19)	Last annotation update)
DE	Colorectal tumor suppressor.	
GN	DCC.	
OS	Rattus norvegicus (Rat).	
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus	
OC	NCBI_TaxId=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=97015074; PubMed=8861902;	
RA	Kelso-Masu K., Masu M., Hink L.,	
RA	Ciullo J.G., Tessier-Lavigne M.;	
RT	"deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";	
RT	Cell 87:175-185(1996).	
RN	[2]	

RE SEQUENCE OF 387-420 FROM N.A.
 RX MEDLINE=90100559; Pubmed=2294591;
 RA Pearson E.R., Cho K.R., Niro J.M., Kern S.E., Simons J.W.,
 RA Ruppert J.M., Hamilton S.R., Preislinger A.C., Thomas G., Kinzler K.W.,
 RA Vogelstein B.;
 RT "Identification of a chromosome 18q gene that is altered in colorectal
 RT cancers.";
 RT Science 247:49-56(1990).
 RL EMBL; U68725; AAB41099.1; -;
 DR EMBL; M32291; AAA41086.1; -;
 DR HSSP; P56376; IYLK.
 DR InterPro; IPR003962; Full repeat.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00041; fn3; 6.
 DR Pfam; PF00047; Ig; 4.
 DR PRINTS; PR00014; ENTYPFII1.
 DR SMART; SM00060; FN3; 6.

Query Match	6.5%;	Score 135;	DB 11;	Length 1445;
Best Local Similarity	25.5%;	Pred. No. 0.0023;		
Matches	56;	Conservative 32;	Mismatches 90;	Total 146

D

QY Db	9 RDPGSGSGNEV-----IEQPQNRVLKSGAENFCTVSQGW--KLIMW 50 : : 215 RNDASTRTGNEAEVRILSDPGLHRLYLFLQRPNSVAIEGCDAYLECCV-GTPPSFTM 273 :: : :
----------	---

QY 51 ASDWADLVSRMEPIITNDEFTQORDOGNFSEIITHNVEPDSGNIRC-SIQNSR 108
Db 274 LRGEERY-----QLRSKKSLTGG-SNLTISNITDDDSGYTCVVTYKKEN 312
QY 109 LHGSAVLVQVNGELFIPSVNLVAENEPCEVTCPSHTWLPDISW-ELGLVSHSSY 167
Db 319 ISASAEVLVLPWPLNHPNSLTYAESMDIEFCASVSGRP-VPTVMKMKNGVVPSPDYF 377
QY 168 FYRPSDIDQASVSLALTPOSGNGLTLCVATWKLARKSA 207
Db 378 QIVGSGNLR----ILGVKCDDEGYQCVVNAENEGNAQSSA 413

```

RESULT 7
Q91562
ID Q91562
AC Q91562;
PRELIMINARY;
PRT; 1427 AA

```

DPT 01-NOV-1996 (TREMBL|rel. 01, created)
 DPT 01-NOV-1996 (TREMBL|rel. 01, last sequence update)
 DDI 01-Dec-2001 (TREMBL|rel. 19, last annotation update)
 Tumor suppressor.
 XDCOA.
 Xenopus laevis (African clawed frog).
 OSC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 CC Xenopodinae; Xenopus.
 NCBI_TaxID=8335;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=95113183; PubMed=7813784;
 FeatCell W.E., Reale M.A., Candia A.F., Wright C.V., Cho K.R.,
 VAA Featon E.R.;
 "Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the nervous system of developing *Xenopus* embryos.";
 TTT Dev. Biol. 166:654-663(1994).
 FMR. rt00064

```

DR  HSSP; 140786; AAAV018v.1; -.
DR  InterPro; IPR003962; FNIII_repeat.
DR  InterPro; IPR003961; FN_III.
DR  InterPro; IPR003598; Ig_C2.
DR  InterPro; IPR003600; Ig_Like.
DR  InterPro; IPR003006; Ig_MHC.
DR  Pfam; PF00041; fn3; 6.
DR  Pfam; PF00047; Ig_4.
DR  PRINTS; PR00014; FN1YPE111.
DR  SMART; SM00060; FN3; 6.
DR  SMART; SM00408; IGC2; 3.
DR  SMART; SM00410; Ig_Like; 2.
KW  Immunoglobulin domain; Repeat.
SQ  SOURCE 1427 AA; 15653 MW; 6169132205770000

```

Query Match	6.4%;	Score 134.5;	DB 13;	Length 1427;
Best Local Similarity	25.4%;	Pred. No. 0.0025;		
Matches	65;	Conservative 33;	Mismatches 93;	Indels 65;

QY 12 PGSGS-GNEV-----LEGPONRVKLSGSAENCTVSOGW--KLIMAL 52
 ||| ||| :||| :||| :||| :|||
 Db 217 PGSAFVGNAELRLISESGILRQOVFLQRNSVAAYEGDALECAVS-GYPTPIVMQ 275
 QY 53 SMMVLSTRMPEPILTNDRETSQRYDGGNFTSEMIINHVPSDSGNIRC--SLQNSRLH 110

Db 276 GD-----EPYPIKTR-----KYSVLGG--SNLLISNTDDACATYCAATYKNENTS 320
 QY 111 GSAYLTVQVGMELPSPNVLVAENPECEVTCL-----PSHWTLPDISW-ELGLVSHSS 165
 Db 321 FSADLTWVPPQFLNHPANLAYESMDIEFECVAGSKPS-----PYVKKTKNEEVIPSD 375
 QY 166 YFVPEPDLQSNVSLTALTPQSNGLTLCVA-----TWKSLKARSAVNLTVI-RC 216
 Db 376 YFQIYDGSNLR-----ILGLVKSDEGYQCIENAGNIQYAOQLIIPDPAPVSSILPSA 431
 QY 217 PDGTGGGINIPGVLS 232
 Db 432 PRDV-----VPIVLS 442

RESULT 8

061987 PRELIMINARY; PRT; 871 AA.
 AC 061987;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Nsk2 protein precursor.
 GN MUSK OR NSK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-MYOBLAST;
 RC MEDLINE-9534951; PubMed-7624144;
 RA Ganju P., Wallis E., Brennan J., Reith A.D.;
 RT "Cloning and developmental expression of Nsk2, a novel receptor
 tyrosine kinase implicated in skeletal myogenesis..";
 RL Oncogene 11:281-290(1995).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: X86444; CAA60165.1; -
 DR HSSP: P11362; IFGK.
 DR MGI: 103581; Musk.
 DR InterPro: IPR002453; Beta.tubulin.
 DR InterPro: IPR000719; Euk.pkinase.
 DR InterPro: IPR000024; Fz_domain.
 DR InterPro: IPR003596; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR002290; Ser_thr.pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF01392; Fz_1.
 DR Pfam: PF00047; Ig_3.
 DR Pfam: PF00069; Pkinase_1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk.pkinase; 1.
 DR SMART: SM00408; Igc2; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS50038; Fz_1.
 DR PROSITE: PS50107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 DR ATP-binding; Immunoglobulin domain; Serine/threonine-protein kinase;
 KW Signal; Transferase.

FT SIGNAL 1
 SQ SEQUENCE 871 AA: 97047 MW; F3C53DC6AF702AB CRC64;

Query Match 6.4%; Score 133.5; DB 11; Length 871;
 Best Local Similarity 21.6%; Pred. NO. 0.0016;
 Matches 73; Conservative 45; Mismatches 127; Indels 93; Gaps 17;

QY 4 GAMENRDPGSGSGNEVIEGPQNAVNLKSGARFNCYISQGWK-LIMALLSDMVVLSVRP 62

Db 113 CALDYKMKP-----KTRRPINVKITEGLKAVLPCTMGNGKPSYSWIKGD----- 158.
 QY 63 MEPIITNDRTFSORYDGGNGFTSEMIHNVPEPDSGNICSLNSLHGSAY-----LTV 117
 Db 159 -NALRENSRIAALE-----SGSLRLHNQKEDAOYRCVAAVNSL--GTANISKLKLEV 208
 QY 118 QVMEELPSPNVLVAENPECEVTC-----PSHWTLPDISW-ELGLVSHSSYFVPEP 172
 Db 209 EVLGRILAPESHVNTGSPVTLACTEIGLP-----VPTISWLENGNNAVSSGIGESVAD 263
 QY 173 SDLOSASVSLTALTPQSNGLTLCVAT-----WKSLSKARSAVNLTVIR-----CPQD 219
 Db 264 RVIDSRLLFLITKP-----GLYTCIATNKHGEKFTYAKAAATVSAEWSKSQDSQGYCAQY 320
 QY 220 TGGGINITG-----VLSSPLSGSLP-----TWKVGGLAGLMTLTPCTLT 264
 Db 321 REGVILMOGPGKMLVLPPTSHRDPEDAOELLIRHNMEL-----KAVSPLCRPAA 373
 QY 265 RCCCCRRCCCGCN-----CCCRCC-----FCGR 287
 Db 374 EALLCYHLFLCSPGVVPTPMPICREYCLAVKELFCAK 411

RESULT 9

061988 PRELIMINARY; PRT; 881 AA.
 AC 061988;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Nsk2 protein precursor.
 GN MUSK OR NSK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-MYOBLAST;
 RC MEDLINE-9534951; PubMed-7624144;
 RA Ganju P., Wallis E., Brennan J., Reith A.D.;
 RT "Cloning and developmental expression of Nsk2, a novel receptor
 tyrosine kinase implicated in skeletal myogenesis..";
 RL Oncogene 11:281-290(1995).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: X86445; CAA60166.1; -
 DR HSSP: P11362; IFGK.
 DR MGI: 103581; Musk.
 DR InterPro: IPR002453; Beta.tubulin.
 DR InterPro: IPR000719; Euk.pkinase.
 DR InterPro: IPR000024; Fz_domain.
 DR InterPro: IPR003596; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR002290; Ser_thr.pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF01392; Fz_1.
 DR Pfam: PF00047; Ig_3.
 DR Pfam: PF00069; Pkinase_1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk.pkinase; 1.
 DR SMART: SM00408; Igc2; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS50038; Fz_1.
 DR PROSITE: PS50107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 DR ATP-binding; Immunoglobulin domain; Serine/threonine-protein kinase;
 KW Signal; Transferase.

FT SIGNAL 1 21 POTENTIAL.
 SQ SEQUENCE 881 AA; 98435 MW; EA0D0282EB28ED7 CRC64;
 Query Match 6.4%; Score 133.5; DB 11; Length 881;
 Best Local Similarity 21.6%; Pred. No. 0.0016;
 Matches 73; Conservative 45; Mismatches 127; Indels 93; Gaps 17;

QY 4 GAMEDNDPPGSGSGNEVEIEGPNARVYKSGQARNCVSGQWK-LIMMALSDMVLVSRP 62
 DB 113 GALQVKKMP-----KITRPINVKIIEGLKAVLPCTTMGNPKPSVSWIKGD----- 158
 QY 63 MEPIITNDPFTSORYDQGNFTSEMIINHPVPSDGNIRCSLONSRLHGSAY-LTV 117
 DB 159 -NALRENSRIALP-----SGSLRIHNVOKEDAGQRCVAKNSL--GTAYSKVLVLEV 208
 QY 118 QVNGELFIPSVNLVVAENPCEVTC-----LPSSHMTWLPDISM-ELGLVSHSSYFVPEP 172
 DB 209 EYLGRIILRAPESHNVTFGSEFVTLCTEIGIP-----VPTIWIENGNAVSSGSIOESVKD 263
 QY 173 SDLOSANSIALTPQSNGLTVCAT-----WKSIAKRSATVNLTVIR-----CPQD 219
 DB 264 RVIDSRQLQLEITKP---GLYTCLATNKHGKEFSTAKAAATVSIAMSKSQKDSQGYCAQY 320
 QY 220 TGGGINIFG-----VLSSLPISLGSFSL-----TWGKVGGLAGTMLTPTCTLT 264
 DB 321 RGEVCMNAVIAKDALVFLNISTYADPEEAQELVHTANNEI-----KVSPVCRPAEAL 373
 QY 265 RCCCCRRCCGCN-----CCCRCC-----FCCR 287
 DB 374 EALCTYHLFLRCSPGVVPFMPICREYCLAVKELFCAR 411

RESULT 10
 ID 015146 PRELIMINARY; PRT; 869 AA.
 AC 015146;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE Muscle specific tyrosine kinase receptor.
 GN MUSK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009854; PubMed=7546737;
 RA Valenzuela D.M., Stitt T.N., Distefano P.S., Rojas E., Mattsson K.,
 RA Compton D.L., Nunez L., Park J.S., Stark J.L., Giles D.R., Thomas S.,
 RA LeBeau M.M., Fernald A.A., Copeland N.G., Jenkins N.A., Burden S.J.,
 RA Glass D.J., Yancopoulos G.D.;
 RT "Receptor tyrosine kinase specific for the skeletal muscle lineage:
 RT expression in embryonic muscle, at the neuromuscular junction, and
 RT after injury.";
 RL Neuron 15:573-584(1995);
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Valenzuela D.M., Rojas E., Yancopoulos G.D.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF006464; AAB63044.1; -
 DR HSSP; P1362; IFGK.
 DR InterPro; IPR002453; Beta_tubulin.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000024; Fz_domain.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00392; Fz; 1.
 DR Pfam; PF00047; Ig; 3.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00038; Fz; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
 KW ATP-binding; Immunoglobulin domain; kinase; Receptor; Transferase;
 KW Tyrosine-protein kinase.
 SQ SEQUENCE 869 AA; 97056 MW; 3DDC20E179FA010C CRC64;
 Query Match 6.3%; Score 132.5; DB 4; Length 869;
 Best Local Similarity 20.8%; Pred. No. 0.002;
 Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;

QY 4 GAMEDNDPPGSGSGNEVEIEGPNARVYKSGQARNCVSGQWK-LIMMALSDMVLVSRP 62
 DB 113 GALQVKKMP-----KITRPINVKIIEGLKAVLPCTTMGNPKPSVSWIKGD----- 158
 QY 63 MEPIITNDPFTSORYDQGNFTSEMIINHPVPSDGNIRCSLONSRLHGSAY-LTVQVNG 121
 DB 159 -SPLENSRIALP-----SGSLRIHNVOKEDAGQRCVAKNSL--GTAYSKVLVLEV 208
 QY 122 ELFIPSVNLVVAENP-----CEVTCPLRSHMTWLPDISM-ELGLVSHSSYFVPEP 172
 DB 209 EYFARILRAPESHNVTFGSEFVTLCTATGIP-----VPTIWIENGNAVSSGSIOESVKD 263
 QY 173 SDLOSANSIALTPQSNGLTVCAT-----WKSIAKRSATVNLTVIRCPDQGGG----- 223
 DB 264 RVIDSRQLQLEITKP---GLYTCLATNKHGKEFSTAKAAATVSIAMSKSQKDNKGYCAQY 320
 QY 224 ---INIPGVLLSLPSLGSFSL-----PTGKVGGLAGTMLTPTCTLTIRCC 267
 DB 321 RGEVCMNAVIAKDALVFLNISTYADPEEAQELVHTANNEI-----KVSPVCRPAEAL 373
 QY 268 CCRRCRCCGCN-----CCCRCC-----FCCRKRGRFIOFQKSKSEKTNKETETE 312
 DB 374 LCNHIEFQECSPGVVPPPIPCREYCLAVKELFCAR-----WLYMEKTHRGLYSEHMLL 429
 QY 313 SGNENSGYNSDEKTDITDASLPKSCSSDPEQRNSSCGPPHQRADQP---PRPASHP 368
 DB 430 SVPECKSLKSMHMDPTACRLP---HLDYKNKEMLKTFPP--MTSKPSVDIPLPSSS 482
 QY 369 QASEFNLA 375
 DB 483 SSSFSVS 489

RESULT 11
 ID Q22048 PRELIMINARY; PRT; 164 AA.
 AC Q22048;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE T0187.8 protein.
 GN T0187.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sins M.A.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z66499; CAA91301.1; -

DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR001271; Defensin_mammal.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001007; VWF_C.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE: PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01208; VWF; UNKNOWN_1.
 SQ SEQUENCE 164 AA; 16499 MW; C002D4BD36C9CED CRC64;

Query Match 6.2%; Score 129.5; DB 5; Length 164;
 Best Local Similarity 32.1%; Pred. No. 0.00043;
 Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;

QY 157 LGLLVSHSYFVPEPSDLSGAVSIALTPQSNGLTCVATWSLKARKSATVNLVYIRC 216
 Db 6 LAIIAIGTFIAV---SQVQSAV-----LPVSSTELATVGTDTASTAIDTGNSSSRV 57
 QY 217 PDVGGGINIPGVLSLPSLGSFLPTWKGVLGAGTMLTPCTLTIRCCCRRCGCG 276
 Db 58 KRGGGCGCGCGCGC-----GCCCGCGGGGG--CGCCCRPRCCCGCCRCCTC 101
 QY 277 --NCCC-RCFCPCR 287
 Db 102 CRCCCTRCCTCR 115

RESULT 12
 Q17641 PRELIMINARY; PRT; 197 AA.

ID 017641
 AC 017641
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 20.6 kDa protein.
 GN C0466.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX "The sequence of C. elegans cosmid C0466.";
 RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX "Direct Submission.";
 RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U55854; AAK68151.1;
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR001271; Defensin_mammal.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001007; VWF_C.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE: PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01208; VWF; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 20596 MW; FBSF9457BFB98AD CRC64;

Query Match 6.1%; Score 127; DB 5; Length 197;

Best Local Similarity 29.6%; Pred. No. 0.00093;
 Matches 32; Conservative 8; Mismatches 42; Indels 26; Gaps 4;

QY 260 CULTRCCCRRCRCGC--NCCC-RCFCPCR-----KGFRIQOK 298
 Db 87 CCRPRCCCRRCRCCTCCCTCCCTCCRCRCGCGCGCGCGGGHFLKSTMAPK 146
 QY 299 KSEKTKETETESGNGSNGSDQKTTDPAVLPKSCSSDPEOR 346
 Db 147 PLIEKAKQENDEKKNRASN-----CCDLASVVRPNPDADNYR 189

RESULT 13
 Q950Y1 PRELIMINARY; PRT; 166 AA.

ID 0950Y1
 AC 0950Y1
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 17.0 kDa protein.
 GN C0466.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX "The sequence of C. elegans cosmid C0466.";
 RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX "Direct Submission.";
 RT Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U55854; AAK68151.1;
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR001271; Defensin_mammal.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001007; VWF_C.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE: PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01208; VWF; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 166 AA; 16971 MW; 9D9D130351BB50F1 CRC64;

Query Match 6.0%; Score 125.5; DB 5; Length 166;
 Best Local Similarity 64.5%; Pred. No. 0.001;
 Matches 20; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 260 CULTRCCCRRCRCGC--NCCC-RCFCPCR 287
 Db 86 CCRPRCCCRRCRCCTCCCTCCCTCCRCRCGCGCGCGGGHFLKSTMAPK 116

RESULT 14
 Q9DF61 PRELIMINARY; PRT; 344 AA.
 ID 09DF61
 AC 09DF61
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE OBCAM alpha 1 isoform.
GN OBCAM.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=2049204; PubMed=11042360;
RA Lodge A.P., Howard M.R., McNamee C.J., Moss D.J.;
RT "Co-localisation, heterophilic interactions and regulated expression
of IGLON family proteins in the chick nervous system.";
RL Brain Res. Mol. Brain Res. 82:84-94(2000).
DR EMBL: AF292934; AAG01871.1; -
DR InterPro: IPR003599; IG.
DR InterPro: IPR003598; IG.
DR InterPro: IPR003600; IG.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; IG_3.
DR SMART: SM00409; IG_3.
DR SMART: SM00408; IG2; 3.
DR SMART: SM00410; IG1like; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 344 AA; 37531 MW; 37FE6051CBF0E7B4 CRC64;
Query Match 6.0%; Score 125; DB 13; Length 344;
Best Local Similarity 25.4%; Pred. No. 0.0029;
Matches 64; Conservative 33; Mismatches 97; Indels 58; Gaps 14;
QY 26 NARYKSGQARNCYVSGWKLIMALSMDVYL-----SVRPEPIITNDRTSQRD 78
DB 44 NTVKGGSATLRCTVDVRRVAV-LNRSTLYAGNDKMSIDNRVILSN---TKQY- 98
QY 79 QGNTSEMIHNPVSDSGNRCSLQ-----NSRLHGSYLTVQWGEFISVNLV 132
DB 99 -----SIKIHNDVYDEGYTCSVQIDNHPKTSRVA---LIVQPPQIVNISSDITV 147
QY 133 AENECEYTCI-----PSHWTLPPDISWELGLVSHSYFVPEPSDQSAVSIATLPQS 188
DB 148 NEGSSVTLMCLAFGRPE-----PYVTWR---HLGKGGGFSEDEYLE---ITGITREQ 195
QY 189 NGLTLCVATWKSILKARKATVNLV-----TRCQDPIGGGGINIPGVL---SSLPSLGF 239
DB 196 SGEVECSAV-NDVAVPDVRAKYVTVNIPYISNAKNTGASVGGKILQCEASAVPAVEFO 254
QY 240 LPTWGXVGLGIA 251
DB 255 --WFKEDTRLA 263

RESULT 15
Q9DBP0 PRELIMINARY; PRT; 697 AA.
AC Q9DBP0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Solute carrier family 34 (sodium phosphate), member 2.
GN SLC34A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carrini P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK004832; BAB23600.1; -
DR MGI: 1342284; SLC34a2.
DR InterPro: IPR003841; Na_P1_cotrans.
DR Pfam: PF02690; Na_P1_cotrans; 1.
DR TIGRfam: TIGR01013; 2a38; 1.
SQ SEQUENCE 697 AA; 76244 MW; 2A7B9384857E16F CRC64;
Query Match 6.0%; Score 125; DB 11; Length 697;
Best Local Similarity 25.8%; Pred. No. 0.0074;
Matches 41; Conservative 12; Mismatches 45; Indels 61; Gaps 8;
QY 233 LPSLGFSLPTWGXVGLGAGTMLTPTGTLTRCCRRRCG-----CNC-CCRCFCG 286
DB 593 LPLMWSLKPMDV-----ISLATTC-FQRRCCCCRCVCCVCGCCRCRSCG 644
QY 287 RRRKRFRIQKKSEKTEKTEKTESGNGNSYNDQKTDYPSLPRKSCSSDPEDR 346
DB 645 R-----DQGE-----EKEEDD---IPKASGATDNAM 671
QY 347 NSSCGPRHORAPRPPASHPQASFNLASPEKVSNTV 385
DB 672 SKEC-----QDEKGVLEVLSMKALSTTV 696

Search completed: April 28, 2003, 21:07:38
Job time : 41.1283 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:24:30 ; Search time 10.3692 Seconds
(Without alignments)
1543.990 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088
Sequence: 1 MVAGAMENRDPGSGSGNENY.....HPQAFNLASPERKSVNTTV 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145.5	7.0	404	1 RAGE_HUMAN	Q15109 homo sapien
2	136	6.5	1447	1 DCC_HUMAN	P43146 homo sapien
3	132	6.3	1447	1 DCC_MOUSE	P70211 mus musculu
4	126	6.0	1914	1 KML5_HUMAN	Q15746 homo sapien
5	125	6.0	337	1 G55A_CHICK	Q98892 gallus gall
6	122	5.8	353	1 CEPV_CHICK	Q90773 gallus gall
7	121	5.8	344	1 NTR1_RAT	Q62718 rattus norv
8	119	5.7	1070	1 PTK7_HUMAN	Q13308 homo sapien
9	114.5	5.5	862	1 CD22_MOUSE	P35329 mus musculu
10	113	5.4	1461	1 NEOL_HUMAN	Q92859 homo sapien
11	112.5	5.4	345	1 OPCM_HUMAN	Q14982 homo sapien
12	112.5	5.4	416	1 RAGE_BOVIN	Q28173 bos taurus
13	111.5	5.3	345	1 OPCM_BOVIN	P11834 bos taurus
14	111.5	5.3	1443	1 NEOL_CHICK	Q90610 gallus gall
15	111	5.3	620	1 SMP_COTJA	Q92154 coturnix co
16	111	5.3	1377	1 NEOL_RAT	P97603 rattus norv
17	110.5	5.3	345	1 OPCM_RAT	P32736 rattus norv
18	110.5	5.2	4393	1 PGBM_HUMAN	P98160 homo sapien
19	109.5	5.2	249	1 CSP_DROME	Q03751 drosophila
20	109.5	5.2	364	1 DMS2_HUMAN	P20138 homo sapien
21	109	5.2	2481	1 DMS2_CAEEL	Q06561 caenorhabdi
22	106.5	5.1	524	1 BUTY_MOUSE	Q62556 mus musculu
23	106	5.0	1092	1 NCAL2_XENLA	P36335 xenopus lae
24	104	5.0	319	1 A33_HUMAN	Q09195 homo sapien
25	104	5.0	764	1 ICCR_DROME	Q08180 drosophila
26	103.5	5.0	365	1 CGAR_MOUSE	P97792 mus musculu
27	103.5	5.0	3707	1 PGBM_MOUSE	Q05793 mus musculu
28	102.5	4.9	333	1 AMAL_DROME	P15364 drosophila
29	102.5	4.9	879	1 PRPP_RAT	Q62786 rattus norv
30	102.5	4.9	890	1 TYO3_RAT	Q06418 homo sapien
31	102	4.8	365	1 CXAR_HUMAN	P78310 homo sapien
32	100.5	4.7	348	1 KIRO_RAT	Q92018 rattus norv
33	99	4.7	359	1 LACH_DROME	Q24372 drosophila

34	99	4.7	830	1 SREC_HUMAN	Q14162 homo sapien
35	98	4.7	261	1 KLB_RAT	P36374 rattus norv
36	98	4.7	873	1 FAS2_DROME	P34082 drosophila
37	98	4.7	1010	1 CONT_CHICK	P14781 gallus gall
38	97.5	4.7	1449	1 VGL2_CYPMI	P33470 porcine tra
39	97.5	4.7	1906	1 KML5_CHICK	P11799 gallus gall
40	97	4.6	198	1 DUC5_MOUSE	P54101 mus musculu
41	97	4.6	439	1 SYG2_DISOM	P24506 discopyge o
42	96.5	4.6	338	1 LAMP_CHICK	Q98919 gallus gall
43	96.5	4.6	880	1 TYO3_MOUSE	P55144 mus musculu
44	96.5	4.6	880	1 TYO3_RAT	P55146 rattus norv
45	96.5	4.6	1147	1 KML5_RABIT	P29294 oryctolagus

ALIGNMENTS

RESULT 1	ALIGNMENTS
RAGE_HUMAN	STANDARD: PRT: 404 AA.
ID	Q15109; Q15279; Q9Y3R3; Q9H2X7;
AC	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Advanced glycosylation end product-specific receptor precursor
DE	(Receptor for advanced glycosylation end products).
GN	AGER OR RAGE.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RC	TISUDE-Lung;
RX	MEDLINE=92340547; PubMed=1378843;
RA	Nepper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
RA	Elliston K., Stern D., Shaw A.;
RT	"Cloning and expression of a cell surface receptor for advanced
RT	glycosylation end products of proteins."
RL	J. Biol. Chem. 267:14998-15004(1992).
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RC	MEDLINE=95137587; PubMed=7835890;
RA	Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,
RA	Inoko H., Ikemura T.;
RT	"Three genes in the human MHC Class III region near the junction with
RT	the class II: gene for receptor of advanced glycosylation end
RT	products, PBX2 homeobox gene and a notch homolog, human counterpart
RT	of mouse mammary tumor gene int-3."
RL	Genomics 23:408-419(1994).
RN	[3]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RC	Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA	Banta A., Spies T., Hood L.;
RT	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RL	[4]
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.
RC	Abelton M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,
RA	Yamamoto H.;
RT	"Molecular heterogeneity of the receptor for advanced glycation
RT	endproducts."
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN	[5]
RP	SEQUENCE FROM N.A. (ISOFORM 2).
RC	Malherbe P., Richards J., Galliard H., Thompson A., Diener C.,
RA	Schuler A., Huber G.;
RT	"cDNA cloning of a novel secreted isoform of the human Receptor for
RT	Advanced Glycation End products (RAGE) and characterization of cells
RT	co-expressing cell-surface scavenger receptors and Swedish mutant
RT	amyloid precursor protein."
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN	[6]
RP	SEQUENCE FROM N.A. (ISOFORM 1).

CC TISSUE-Lung;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-12 FROM N.A.
 RA Hudson B.I., Puters T.S.;
 RT "Novel polymorphisms in the receptor for advanced glycation
 end products (AGE) gene";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
 PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 RATE IN DIABETES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
 CC -1- Secreted (isoform 2).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/RAESEC;
 CC -1- ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M91211; AAA03574.1; -;
 DR EMBL; D28769; BAA05958.1; -;
 DR EMBL; U89336; AAB47491.1; -;
 DR EMBL; AB036432; BAA89369.1; -;
 DR EMBL; AJ133822; CAB43108.1; -;
 DR EMBL; BC020669; AAR20669.1; -;
 DR EMBL; AF208289; AAG35728.1; -;
 DR Genbank; HGNC:320; AGER.
 DR MIM; 600214; -;
 DR InterPro; IPR003006; Iq_MHC.
 DR InterPro; IPR003598; Iq_C2.
 DR InterPro; IPR003600; Iq_Like.
 DR Pfam; PF00047; Iq_2.
 DR SMART; SM00410; Iq_Like; 1.
 DR SMART; SM00408; Iq_C2; 1.
 DR PROSITE; PS00290; Iq_MHC; 1.
 KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
 KW Alternative splicing; Polymorphism.
 FT CHAIN 1 22
 FT SIGNAL 23 404
 FT DOMAIN 23 342
 FT TRANSMEM 343 363
 FT DOMAIN 364 404
 FT DOMAIN 137 215
 FT DOMAIN 252 308
 FT DISULFID 38 99
 FT DISULFID 144 208
 FT DISULFID 259 301
 FT CARBOHYD 81 384
 FT DOMAIN 380 384
 FT VARSPLIC 54 67
 FT VARSPLIC 275 404
 FT VARIANT 100 100
 FT CONFLICT 1 1
 M -> G (IN REF. 1).
 M -> G (IN REF. 1).

SO SEQUENCE 404 AA: 42802 MW: 0D584C436C30CCE7 CRC64;
 Query Match 7.0%; Score 145.5; DB 1; Length 404;
 Best Local Similarity 23.5%; Pred. NO. 0.00018;
 Matches 77; Conservative 34; Mismatches 94; Indels 123; Gaps 15;
 QY 64 EPIITNDRETS-----ORYDQGNFT--SEMIINHVPSDGNR-----CSLONSRLHGS 112
 DB 162 KPLVPEKEVSVKEQTRRHPEFTGLTLOSLM---VIPARQDPRPTSCSFSPGLRRH 218
 QY 113 ALTYQVMELEIP-----SVNLVAENEP-----CEVTCLPSTMTWLPDIS 154
 DB 219 ALRTAPIQPRWEPEVPLEVOLV---EPGGAAPGVTILCEVPAQS-----PQH 270
 QY 155 WELGLVSHSSYFVPEPSDQSAVSIATLPSSNGTLTCVATMKSKAKSATVNLTY 214
 DB 271 WKMD-----GVPLPLPSPVLLPEIQPODQGYSCVAHSHSGPQESRAVSIIT 321
 QY 215 RQPDQTGGINIPVLSLPSLGFSLPTWKGVLGLAGTMLT-----PCTLTIRCC 267
 DB 322 E-PGEEG-----PTAGSVGGSGGLTALALGILGGLGTRALLGYI 361
 QY 268 CCRRCGCCNCCRCPCRCRRKRGFRIOPKKSEKKT--NKETPESGNSGYSNDEQ 325
 DB 362 LMQRR-----ORGERKAPENQEEBERAELN----- 389
 QY 326 KTTDTASLPKSCSSDPEQRNSCGPP 353
 DB 390 -----QSEPEAGESESSTGP 404
 RESULT 2
 DCC_HUMAN STANDARD; PRT; 1447 AA.
 ID DCC_HUMAN
 AC P43146;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor suppressor protein DCC precursor (Colorectal cancer suppressor).
 GN DCC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95011532; PubMed-7926722;
 RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
 RT Vogelstein B.;
 RT "The DCC gene product in cellular differentiation and colorectal
 tumorigenesis.";
 RL Genes Dev. 8:1174-1183(1994).
 RN [2]
 RP SEQUENCE OF 1-750 FROM N.A.
 RX MEDLINE-90100559; PubMed-2294591;
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Ruppert J.M.,
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
 RT Vogelstein B.;
 RT "Identification of a chromosome 18q gene that is altered in
 colorectal cancers.";
 RL Science 247:49-56(1990).
 RN [3]
 RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).
 RX MEDLINE-91121517; PubMed-1991322;
 RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
 RA Oliner J.D., Kinzler K.W., Vogelstein B.;
 RT "Scrambled exons.";
 RL Cell 64:607-613(1991).
 RN [4]
 RP GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.
 RX MEDLINE-94245241; PubMed-8188295;
 RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;

[illegible]

RC STRAIN-BALB/c; TISSUE-Brain;
 RL Cooper H.M.;
 CC Submitted (Jun-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
 CC produced by alternative splicing. A third isoform; C is
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
 CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
 CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
 CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
 CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION
 CC AND REMAIN AT THIS LEVEL IN THE ADULT.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X85788; CAA59786.1; -
 DR HSSP; P56276; 1TLK.
 DR MGD; MGI:94869; DCC.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; Ig; 4.
 DR PRINTS: PR00014; PNTYPEP11.
 DR SMART; SM00060; FN3; 6.
 DR SMART; SM00410; Ig_Like; 2.
 DR SMART; SM00408; IgC2; 3.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
 FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT INT_MET 85 1447
 FT DOMAIN 26 1097
 FT TRANSMEM 1098 1122
 FT DOMAIN 1123 1447
 FT DOMAIN 54 124
 FT DOMAIN 154 219
 FT DOMAIN 254 317
 FT DOMAIN 345 407
 FT DOMAIN 426 522
 FT DOMAIN 525 618
 FT DOMAIN 619 716
 FT DOMAIN 722 816
 FT DOMAIN 840 940
 FT DOMAIN 941 1042
 FT DISULFID 61 117
 FT DISULFID 161 117
 FT DISULFID 261 310
 FT DISULFID 352 400
 FT CARBOHYD 60 94
 FT CARBOHYD 94 94
 FT CARBOHYD 299 299
 FT CARBOHYD 318 318
 FT CARBOHYD 478 478
 FT CARBOHYD 628 628
 FT CARBOHYD 702 702
 FT VARSPLIC 819 838

SEQUENCE 1447 AA; 158298 MW; 0DLE1097C22D5B9F CRC64;
 Query Match 6.3%; Score 132; DB 1; Length 1447;
 Best Local Similarity 25.3%; Pred. No. 0.011;
 Matches 56; Conservative 31; Mismatches 91; Indels 42; Gaps 9;
 9 RPPPSGSGNEV-----IEGPOARVLKSGQAFNCTVQGM--KLIM 50
 215 RNPASIRIGNEAEVRLILSDPGLHQVLFQRPNSVIAIEGDAVLECYG-GRPPSFTW 273
 51 ALSDWVTVSRPMEITINDFTSQRDQNGFTSEMIHNVPSDSGNIRC--SLQNSR 108
 274 LRGEVIT-----OLRSKYSILGG--SNLLISNVTDSDGTYTCVYTKNEN 318
 109 LHSAYLVTVQVGELEFIPSVNLVVAENPEVCIPSHWFWLPDISW-ELGLVHSSTY 167
 319 ISASAEFLVLPWFELNHPNSLVAYESMDIEFCANVSGK-VFTVMKNGDVVLPISDYF 377
 168 FVPEPSDLSAVSLALTPQSGTITCTVATWKSLSKARKSA 207
 378 QVGSNLR-----ILGVYKSDGEFYQCVAEENAGNAQSSA 413
 Db
 RESULT 4
 KMLS_HUMAN STANDARD; PRT: 1914 AA.
 AC Q15746; Q14844; Q16794; Q95796; Q95797; Q95798; Q95799; Q9UBG5;
 AC Q9UBT9;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin light chain kinase, smooth muscle and non-muscle isozymes
 DE (EC 2.7.1.117) (MCK) [contains: Telokin (kinase related protein)
 DE (KRP)].
 GN MYLK OR MLCK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Birukov K.G., Garcia J.G.N.;
 RC TISSUE=umbilical vein endothelial cells;
 RX MEDLINE=97304466; PubMed=9160829;
 RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,
 RA Verin A.D.;
 RA "Myosin light chain kinase in endothelium: molecular cloning and
 RA regulation.";
 RA Am. J. Respir. Cell Mol. Biol. 16:489-494(1997).
 RN [2]
 RP REVISIONS.
 RA Birukov K.G., Garcia J.G.N.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).
 RA TISSUE=umbilical vein;
 RX MEDLINE=99216419; PubMed=10198165;
 RA Lazar V.L., Garcia J.G.N.;
 RA "A single human myosin light chain kinase gene (MLCK; MYLK).";
 RA Genomics 57:256-267(1999).
 RN [4]
 RP REVISIONS (ISOFORM 2).
 RA Birukov K.G., Garcia J.G.N.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 923-1914 FROM N.A.
 RA TISSUE=hippocampus;
 RX MEDLINE=96121365; PubMed=8575746;
 RA Potier M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossier J.,
 RA Turnell W.G.;
 RA "The human myosin light chain kinase (MLCK) from hippocampus:
 RA cloning, sequencing, expression, and localization to 3gen-921.";
 RA Genomics 29:562-570(1995).
 RN [6]

RESULT 5	G55A_CHICK		
ID	G55A_CHICK	STANDARD;	PRT; 337 AA.
AC	096892;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Neurite inhibitor GP55-A precursor (OBCCAM protein gamma isoform).		
OS	Gallus gallus (chicken).		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
CC	Gallus.		
OX	NCBI_TaxID=9031;		
ON	11		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=Brain;		
RA	Lodge A.P., Kim D.-S., Howard M.R., McNamee C.J., Smith N., Moss D.J.,		
RT	"Cloning of CEPU-s, a secreted isoform of CEPU-1, and OBCCAM cDNAs from		
RT	chick: structural diversity of Igron family proteins."		
RL	Submitted (APR-1999) to the EMBL/genbank/DDJ databases.		
RI	[2]		
RP	SEQUENCE OF 64-337 FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=Brain;		
RX	MEDLINE=97157768; PubMed=9004047;		
RA	Wilson D.J.A., Kim D.-S., Clarke G.A., Marshall-Clarke S., Moss D.J.,		
RT	"A family of glycoproteins (GP55), which inhibit neurite outgrowth,		
RT	are members of the Ig superfamily and are related to OBCCAM,		
RT	neurotrophin, LAMP and CEPU-1."		
J	J. Cell Sci. 109:3129-3138(1996)		

```

0Y 26 NARVLKSOARENCYTSQCKLIMALSDMYVL-----SVRMEPIITTDRTSQRD 78
Db 37 NVTVROSESATLCTCAYDDRERRVAM-LNRSTLYAGDNKMSIDNRVILSN---TKIY- 91
0Y 79 QGQNFTESEMIHNWEPSDSGNIRCSLO-----NSRHSAGYLTVQYMELETSVNLV 13
Db 92 -----SIHNVDYDIDGPIITCSVQIDNHPKTSRVH-----LTVQVPOIVNISSDITV 14
0Y 133 AENEPCCVYCL-----PSHWTLIPDISWELGLVSHSSSYFVPEPSDQASVILALTPOS 188
Db 141 NEGSSTYIMLCLAGRE-----PVTWR-----HLSGKQGFVSEDTLE-----ITGIRFEQ 188
0Y 189 NGTLTCAVATKSLKARKSATVNLTV-----IRCPQDGGGINIPVL-----SLSPJSGFS 239
Db 189 SGEECSAAV-NDYAVVDKRAKVIVNIPPTISNAKNKGASVQCGKILQCEASAPVAFEQ 247
0Y 240 LPTMGKVGIGLA 251
Db 248 ---WFKEDTRLA 256

```

RESULT 6		
CEPU_CHICK	STANDARD;	PRT; 353 AA.
ID CEPU_CHICK		
AC Q90773;		
DT 01-NOV-1997 (Rel. 35)		Created)
DT 01-NOV-1997 (Rel. 35)		Last sequence update)
DT 15-JUN-2002 (Rel. 41)		Last annotation update)

DE CEPV-1 protein precursor.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_Taxid=9031;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RC MEDLINE=96370549; PubMed=8774445;
 RA Spatmann F., Brumentorf T.;
 RT CEPV-1, a novel immunoglobulin superfamily molecule, is expressed by
 RL developing cerebellar Purkinje cells.";
 CC J. Neurosci. 16:1770-1779(1996).
 CC -1- FUNCTION: IT MAY BE A CELLULAR ADDRESS MOLECULE SPECIFIC TO
 CC PURKINJE CELLS. IT MAY REPRESENT A RECEPTOR OR A SUBUNIT OF A
 CC RECEPTOR COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC (SHOWN HERE) ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: FOUND ON THE DENDRITES, SOMATA AND AXONS OF
 CC DEVELOPING PURKINJE CELLS. UNDETECTABLE ON OTHER NEURONS LIKE
 CC GOLI OR GRANULE CELLS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BY DEVELOPING CEREBELLAR PURKINJE
 CC CELLS. EXPRESSION COINCIDES WITH THE GROWTH OF THE DENDRITIC TREE.
 CC AFTER PURKINJE CELLS HAVE FINISHED THEIR MIGRATION FROM THE
 CC VENTRICULAR ZONE (FROM E15 UNTIL E21). EXPRESSED IN THE ADULT
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z72497; CAA96578.1; -
 CC InterPro: IPR003006; Iq_MHC.
 CC InterPro: IPR003598; Iq_C2.
 CC InterPro: IPR003600; Iq_Like.
 CC InterPro: PF00047; Iq_3
 CC Pfam: PF00047; Iq_3
 CC SMART: SM00410; Iq_Like; 1.
 CC DR SMART; SM00408; IGC2; 2.
 CC DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 CC Repeat; Signal; Alternative splicing.
 CC KW SIGNAL.
 CC FT CHAIN 1 28 POTENTIAL.
 CC FT PROPEP 29 330 CEPV-1 PROTEIN.
 CC FT DOMAIN 331 353 REMOVED IN MATURE FORM (POTENTIAL).
 CC FT DOMAIN 48 120 IG-LIKE C2-TYPE DOMAIN 1.
 CC FT DOMAIN 148 206 IG-LIKE C2-TYPE DOMAIN 2.
 CC FT DOMAIN 234 300 IG-LIKE C2-TYPE DOMAIN 3.
 CC FT DISULFID 55 113 POTENTIAL.
 CC FT DISULFID 155 199 POTENTIAL.
 CC FT DISULFID 241 293 POTENTIAL.
 CC FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT LIPID 330 330 GPI-ANCHOR (POTENTIAL).
 CC FT VARSPLIC 310 320 MISSING (IN MAJOR ISOFORM).
 CC SQ SEQUENCE 353 AA; 38736 MW; 2550C48591EBBBA6 CRC64;
 CC
 CC Query Match 5.8%; Score 122; DB 1; Length 353;
 CC Best Local Similarity 25.3%; Pred. No. 0.012;
 CC Matches 62; Conservative 36; Mismatches 91; Indels 56; Gaps 13;
 CC 26 NARYIGSQARFNCYVSGW---KLIMWALSDWVLSVRPEPIITNDRTPTQRY 77

Db 42 NTVIGGSAATLRCSVDNRTVRANLRSSILYAGNDKWCIDPRVLLANTKTQYSIQ-- 99
 Oy 78 DOGNGFTSEMIHNVEPSDSGNIRCSIQ-----NSRLGSAVLYQVNGELFIPSNVY 131
 Db 100 -----IHVDVYDEGPTCSQVQDNHPRKTSRVH---LIVQSPKITESSSDS 144
 Oy 132 VAENPECEVTCPLSHMTMLPD--ISWELGLVSHSSVYEPSPDQSAVSLIATLPQSN 189
 Db 145 INEGQNSVLTCTIA---IGRPDPTITWR---HISPKAVGFISEDEYLE-----ITGITREOS 194
 Oy 190 GLTICVATWKSILKARKSATVNLVY-----IRCPDQTGGGINITGVY-----SSLPISGFSL 240
 Db 195 GYEESAS--NVAAPVQVRKATVNPYPYISDAKSTGVPGQKGLMCEASAVPSDFQ- 252
 Oy 241 PWWGK 245
 Db 253 --WYK 255
 RESULT 7
 ID NTPL_RAT STANDARD; PRT; 344 AA.
 AC 062718;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE Neurotrophin precursor (GP65).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_Taxid=10116;
 CC (1)
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.
 RP SPRAIN-Sprague-Dawley;
 RC MEDLINE=95198094; PubMed=7891157;
 RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Enstachio P.,
 RA Salzer J.L.;
 RT Cloning of neurotrophin defines a new subfamily of differentially
 RT expressed neural cell adhesion molecules.";
 RL J. Neurosci. 15:2141-2156(1995).
 CC -1- FUNCTION: NEURAL CELL ADHESION MOLECULE.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN SEVERAL
 CC DEVELOPING PROTECTION SYSTEMS: IN NEURONS OF THE THALAMUS,
 CC SUBPLATE, AND LOWER CORTICAL LAMINAE IN THE FOREBRAIN AND IN THE
 CC PONTINE NUCLEUS, CEREBELLAR GRANULE CELLS, AND PURKINJE CELLS IN
 CC THE HINDRAIN.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: 016845; AAA67445.1; -
 CC InterPro: IPR003006; Iq_MHC.
 CC InterPro: IPR003598; Iq_C2.
 CC InterPro: IPR003600; Iq_Like.
 CC Pfam: PF00047; Iq_3
 CC SMART: SM00410; Iq_Like; 1.
 CC DR SMART; SM00408; IGC2; 2.
 CC DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 CC Repeat; Signal.
 CC KW SIGNAL.
 CC FT CHAIN 1 31 POTENTIAL.
 CC FT PROPEP 32 321 NEUROTRPHIN.
 CC FT DOMAIN 50 122 REMOVED IN MATURE FORM (POTENTIAL).
 CC IG-LIKE C2-TYPE DOMAIN 1.

FT DOMAIN 150 208 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 236 302 IG-LIKE C2-TYPE DOMAIN 3.
 FT DISULFID 57 115 POTENTIAL.
 FT DISULFID 157 201 POTENTIAL.
 FT CARBOHYD 243 295 POTENTIAL.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 321 321 GPI-ANCHOR (POTENTIAL).
 SQ SEQUENCE 344 AA: 37998 MW: CBB39BE53B3B224 CRC64;

Query Match 5.8%; Score 121; DB 1; Length 344;
 Best Local Similarity 24.1%; Pred. No. 0.014;
 Matches 59; Conservative 37; Mismatches 93; Indels 56; Gaps 13;

QY 26 NARVAKGSQARNCCTVSGWKLIMVALSDMYLS-----VRPEPIITNDRETSQRYD 79
 DB 44 NNTVQGSATLCTIDIRVTVAMLRSTLYAGNDKMLDPRVYLSN---TQYR-- 98
 QY 80 GGNFTSEMIINVEPSDGNIRCSIQ-----NSRLHGSALVYOVWGELEFISVNLVYA 133
 DB 99 -----SIEIGNVDYDEGPYTCVQVDNHPKTSRVA---LTVQSPKIVEISSDISIN 148
 QY 134 ENPECEVLC-----PSHWMLPDISWELGLVSHSYFVEPEPSDQASVIALPOSN 189
 DB 149 EGNNTSLCTIATVGRPE-----PVTWR---HISPKAVGFSEDEYLE-----IQGTRRS 196
 QY 190 GTLLCVATWKSARKSATVNLTVIRP-----QDGGGGINIPGVL-----SSPLSLGFSL 240
 DB 197 GEYECSSAS-NDVAPVYRNVNVPPIYSKAGTGVPGKGLQCEASAVPSAERO- 254
 QY 241 PTWCK 245
 DB 255 --WFK 257

RESULT 8
 PKT7_HUMAN STANDARD; PRT; 1070 AA.
 AC 013308: 013417;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase-like 7 precursor (colon carcinoma kinase-4) (CCR-4).
 GN PKT7 OR CCK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-Colon carcinoma, and Placenta;
 RX MEDLINE=96074849; PubMed=7478540;
 RA Mossie K., Jallal B., Alves F., Sures I., Plozman G.D., Ullrich A.;
 RT "Colon carcinoma kinase-4 defines a new subclass of the receptor tyrosine kinase family.";
 RL Oncogene 11:2179-2184(1995).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE-Fibroblast;
 RX MEDLINE=97037064; PubMed=8882711;
 RA Park S.-K., Lee H.-S., Lee S.-T.;
 RT "Characterization of the human full-length PKT7 cDNA encoding a receptor protein tyrosine kinase-like molecule closely related to chick KLG.";
 RL J. Biochem. 119:235-239(1996).
 [3]

RP SEQUENCE FROM N.A., AND REVISION TO 834.
 RA Lee S.-T., Park S.-K., Lee H.-S., Ji A.R., Jung J.W.;
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR PROGRESSION MARKER.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN RECEPTOR SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: U33635; AAC87565.1; -
 CC EMBL: U40271; AAC50484.2; -
 CC EMBL: AF447176; AAL39062.1; JOINED.
 CC EMBL: AF447157; AAL39062.1; JOINED.
 CC EMBL: AF447158; AAL39062.1; JOINED.
 CC EMBL: AF447162; AAL39062.1; JOINED.
 CC EMBL: AF447164; AAL39062.1; JOINED.
 CC EMBL: AF447167; AAL39062.1; JOINED.
 CC EMBL: AF447170; AAL39062.1; JOINED.
 CC EMBL: AF447171; AAL39062.1; JOINED.
 CC EMBL: AF447173; AAL39062.1; JOINED.
 CC EMBL: AF447174; AAL39062.1; JOINED.
 CC EMBL: AF447175; AAL39062.1; JOINED.
 CC HSSP: P08631; IAD5.
 CC GeneW: HGNC:9618; PKT7.
 CC MIM: 601890; -
 CC DR InterPro: IPR000719; Euk_pkinase.
 CC DR InterPro: IPR003006; Ig_MHC.
 CC DR InterPro: IPR003598; Ig_C2.
 CC DR InterPro: IPR003600; Ig_Like.
 CC DR InterPro: IPR002011; RTKinase1.
 CC DR InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00047; Ig_7.
 CC Pfam: PF00069; pkinase.1.
 CC PRINTS: PR00109; TYRKINASE.
 CC PRODOM: PD000001; Euk_pkinase.1.
 CC SMART: SM00410; Ig_Like; 2.
 CC SMART: SM00408; IGC2; 5.
 CC SMART: SM00219; TYRK; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00235; RECEPTOR_TYR_KIN_II; FALSE_NEG.
 CC Receptor: Transmembrane; Signal; Glycoprotein; Cell adhesion; Immunoglobulin domain; Repeat.
 CC SIGNAL 1 30
 FT CHAIN 31 1070 TYROSINE-PROTEIN KINASE-LIKE 7.
 FT DOMAIN 31 704 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 705 725 POTENTIAL.
 FT DOMAIN 726 1070 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 46 108 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 143 208 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 239 308 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 336 398 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 426 488 IG-LIKE C2-TYPE DOMAIN 5.
 FT DOMAIN 517 577 IG-LIKE C2-TYPE DOMAIN 6.
 FT DOMAIN 606 671 IG-LIKE C2-TYPE DOMAIN 7.
 FT DOMAIN 796 1066 PROTEIN KINASE. INACTIVE.
 FT DISULFID 53 101 BY SIMILARITY.
 FT DISULFID 150 200 BY SIMILARITY.

FT DISULFID 246 301 BY SIMILARITY.
 FT DISULFID 343 391 BY SIMILARITY.
 FT DISULFID 433 481 BY SIMILARITY.
 FT DISULFID 524 570 BY SIMILARITY.
 FT DISULFID 613 664 BY SIMILARITY.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 92 92 P -> R (IN REF. 2 AND 3).
 FT CONFLICT 147 147 K -> T (IN REF. 2 AND 3).
 FT CONFLICT 207 207 S -> G (IN REF. 2 AND 3).
 FT CONFLICT 495 496 VL -> RV (IN REF. 2 AND 3).
 FT CONFLICT 515 515 G -> E (IN REF. 2 AND 3).
 FT CONFLICT 881 881 E -> G (IN REF. 2 AND 3).
 FT CONFLICT 969 969 A -> P (IN REF. 2 AND 3).
 FT CONFLICT 992 992 S -> F (IN REF. 2 AND 3).
 SQ SEQUENCE 1070 AA: 118260 MM: 47CDF25B8E3698A5 CRC64;
 Query Match 5.78; Score 119; DB 1; Length 1070;
 Best Local Similarity 25.3%; Pred. No. 0.084;
 Matches 58; Conservative 32; Mismatches 103; Indels 36; Gaps 10;
 QY 20 VIEPQNAVYKSGQAFNCVTS-OGWKLIMWALSDVYVSPRPPITITNDFTISQRYD 78
 DB 227 VLAPODYVYKRYEAFNHCFOFSNPPSLQMFEDETPTNNSRPHLRATVFA---- 282
 QY 79 OGNETSEMILHNPSPDSGNIRCSLQNSR-----LHGSATLVYQWAGELFISVNLVYA 133
 DB 283 -----NGSLILLYQVRPNAGYRCIGOGORPPILTEATLHLEIEMDPFEPRVTPAGS 337
 QY 134 ENPEPVTCLPSHTWLPDISME-LGL-LVSHSSYFVPPSPDQSAVSLATLPQSN-G-190
 DB 338 EE---RVTCLEPKGLPEPSVWMEHAGVRLPHGRVY-----QKHELVLANIAESDAG 387
 QY 191 TLTCVATWKSILKARKSATVNLTV-----TRCPDGTGGGINIPGVLSL 233
 DB 388 VYTCIAA--NLAGRRDQVNTVATVPSMLKPKDQSLKEGKPGYLDCL 434
 RESULT 9
 CD22_MOUSE STANDARD: PRT; 862 AA.
 AC P35329;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE B-cell receptor CD22 precursor (Iu-14) (B-lymphocyte cell adhesion molecule) (BL-CAM).
 GN CD22 OR LYB-8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBITaxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DNA/2J, and BALB/c; TISSUE=Liver;
 RX MEDLINE=93315834; PubMed=8100843;
 RA Law C.-L., Torres R.M., Sundberg H.A., Parkhouse R.M., Brannan C.I., Copeland N.G., Jenkins N.A., Clark E.A.;
 RT "Organization of the murine Cd22 locus. Mapping to chromosome 7 and characterization of two alleles."
 RL J. Immunol. 151:175-187(1993).
 CC -1- FUNCTION: MEDIATES B-CELL, B-CELL INTERACTIONS. MAY BE INVOLVED IN THE LOCALIZATION OF B-CELLS IN LYMPHOID TISSUES. BINDS STIMULATED GLYCOPROTEINS: ONE OF WHICH IS CD45.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CD22-ALPHA AND CD22-BETA (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: B-LYMPHOCYTES.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: L16928; AAA02562.1;
 DR MGD; MGI:88322; CQ22.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_6.
 DR SMART: SM00408; IgC2_4.
 DR SMART: SM00410; Ig_Like; 1.
 KW Glycoprotein; Cell adhesion; Transmembrane; Signal; B-cell; Repeat; Immunoglobulin domain; Alternative splicing.
 FT SIGNAL 1 21
 FT CHAIN 22 862
 FT DOMAIN 22 702
 FT TRANSSEM 703 721
 FT DOMAIN 722 862
 FT DOMAIN 158 235
 FT DOMAIN 265 331
 FT DOMAIN 361 418
 FT DOMAIN 450 506
 FT DOMAIN 537 593
 FT DOMAIN 624 681
 FT DISULFID 41 171
 FT DISULFID 46 106
 FT DISULFID 165 229
 FT DISULFID 272 324
 FT DISULFID 368 411
 FT DISULFID 457 499
 FT DISULFID 544 586
 FT DISULFID 631 674
 FT CARBOHYD 105 105
 FT CARBOHYD 116 116
 FT CARBOHYD 139 139
 FT CARBOHYD 168 168
 FT CARBOHYD 265 265
 FT CARBOHYD 275 275
 FT CARBOHYD 378 378
 FT CARBOHYD 408 408
 FT CARBOHYD 460 460
 FT CARBOHYD 561 561
 FT CARBOHYD 589 589
 FT VARIANT 15 15
 FT VARIANT 19 19
 FT VARIANT 76 76
 FT VARIANT 83 83
 FT VARIANT 90 91
 FT VARIANT 94 94
 FT VARIANT 102 102
 FT VARIANT 173 173
 FT VARIANT 179 179
 FT VARIANT 186 186
 FT VARIANT 190 190
 FT VARIANT 192 192
 FT VARIANT 236 236
 FT VARIANT 241 241
 FT VARIANT 244 244
 FT VARIANT 616 616
 S -> I (IN BALB/C).
 R -> H (IN BALB/C).
 E -> K (IN BALB/C).
 P -> S (IN BALB/C).
 V -> I (IN BALB/C).
 Y -> S (IN BALB/C).
 R -> K (IN BALB/C).
 E -> G (IN BALB/C).
 O -> R (IN BALB/C).
 A -> V (IN BALB/C).
 K -> T (IN BALB/C).
 R -> T (IN BALB/C).
 KAEP -> ATTPEDPS (IN BALB/C).
 PP -> LS (IN BALB/C).
 R -> G (IN BALB/C).
 S -> R (IN BALB/C).
 E -> G (IN BALB/C).
 O -> K (IN BALB/C).
 Y -> I (IN BALB/C).
 V -> S (IN BALB/C).
 P -> S (IN BALB/C).
 R -> K (IN BALB/C).
 E -> C (IN BALB/C).
 R -> C (IN BALB/C).
 R -> H (IN BALB/C).
 S -> I (IN BALB/C).

FT VARIANT 714 714 C -> F (IN BALB/C).
 FT VARIANT 787 787 M -> T (IN BALB/C).
 FT VARIANT 808 808 S -> T (IN BALB/C).
 SQ SEQUENCE 862 AA; 96582 MW; A7662D6E87038E83 CRC64;
 Query Match
 Best Local Similarity 19.1%; Score 114.5; DB 1; Length 862;
 Matches 86; Conservative 36; Mismatches 126; Indels 203; Gaps 18;
 QY 10 DPGSGSGNEVIEGPNKRVKSGQARN-----CLVSGQMKL-----TMA 51
 DB 472 NPGSGSS-----VLKPVLRKQKVTWDSMPVSCACNKKSMALPVILANHYA 519
 QY 52 LQDMVYLSRPMPIITNDRTFSQ-----RYDGGNFTSE-----MIHNVPS 95
 DB 520 PDVAVLVKVPASELRAGQAVLQCDFAESNPAERFPMKNGSLVQEGVLSFGSVP 579
 QY 96 DSGNRCISLONSRLHGSAYLVQVWGLFITSVNLV-----VAENEP 138
 DB 580 DSGNRCISLONSRLHGSAYLVQVWGLFITSVNLV-----VAENEP 138
 QY 139 EYTCIP-----SHWTWLPDISWELGLVSHSYFVPEPSDLOSASVILATP---QSN 189
 DB 628 TLSCSDANPPIISQYTW-----FSSQDLHSSQKRLPEPLEVYHT 669
 QY 190 GLTFCVATWKSILKARKSATVNLVIRCPDGTGGGINIPGLVSLPLGFSLPTRWKGVLG 249
 DB 670 GSYRKGT-NGICGESPPSTLVYYSPEITG-----KVALG 706
 QY 250 LAGTMLTPCTTTRCCGCCRRCCGCCGCCRRKRGFRIO-----PKX 298
 DB 707 LGFCLTI--CIIAI-----WGKIKKKKKONRSGQGLQ 738
 QY 299 KSE-----KEKTKETTESGNGNS-----YNSDEKTTDT 330
 DB 739 NSSQSSEFVKKKARRKRIPLSEGPQSGCYNPAMDITVYAILRFPESDMHNMAGTPTAT 798
 QY 331 ASLPKSCS-----SDPEQRNSC 350
 DB 799 QAPPNNSDSVTYSVIOKRPMDGEYVNVNSC 829
 RESULT 10
 NEOL_HUMAN
 ID NEOL_HUMAN STANDARD; PRT; 1461 AA.
 AC Q92859; 000340;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neogenin precursor.
 GN NEOL OR NGN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Fetal brain;
 RC MEDLINE=9723653; PubMed=9121761;
 RA Meyerhardt J.A., Look A.T., Bigner S.H., Fearon E.R.;
 RT "Identification and characterization of neogenin, a DCC-related
 RT gene.";
 RL Oncogene 14:1129-1136(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Fetal brain;
 RC MEDLINE=9731269; PubMed=9169140;
 RA Vielmetter J., Chen X.-N., Miskovich F., Lane R.P., Yamakawa K.,
 RA Korenberg J.R., Dreyer W.J.;
 RT "Molecular characterization of human neogenin, a DCC-related protein,
 RT and the mapping of its gene (NEO1) to chromosomal position 15q22.3-
 RT q23.";
 RL Genomics 41:414-421(1997).

CC CC -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
 CC CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
 CC CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
 CC CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
 CC CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) AND 2;
 CC CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED AND ALSO IN CANCER CELL
 CC CC LINES.
 CC CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC CC SUBFAMILY.
 CC CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL: U61262; AAB17263.1; -
 CC CC EMBL: U72391; AAC51287.1; -
 CC CC HSSP: P02751; 1TTF.
 CC CC Genew: HGNC:7754; NEOL.
 CC CC MIM: 601907;
 CC CC InterPro: IPR003961; FN_III.
 CC CC InterPro: IPR003962; FNIII_repeat.
 CC CC InterPro: IPR003963; IG_MHC.
 CC CC InterPro: IPR003598; IG_C2.
 CC CC InterPro: IPR003600; IG_1like.
 CC CC Pfam: PF00041; fn3; 6.
 CC CC Pfam: PF00047; fn3; 4.
 CC CC PRINTS: PR00014; FNTPETII.
 CC CC SMART: SM00400; FN3; 6.
 CC CC SMART: SM00410; IG_1like; 1.
 CC CC SMART: SM00408; IG_C2; 3.
 CC CC Transmembrane; Immunoglobulin domain; Glycoprotein; Signal; Repeat;
 CC CC Alternative splicing.
 CC CC FT CHAIN 1 33
 CC CC FT DOMAIN 34 1461
 CC CC FT TRANSMEM 34 1105
 CC CC FT DOMAIN 1106 1126
 CC CC FT DOMAIN 1127 1461
 CC CC FT DOMAIN 67 136
 CC CC FT DOMAIN 166 228
 CC CC FT DOMAIN 263 327
 CC CC FT DOMAIN 355 417
 CC CC FT DOMAIN 436 533
 CC CC FT DOMAIN 536 629
 CC CC FT DOMAIN 630 729
 CC CC FT DOMAIN 735 829
 CC CC FT DOMAIN 850 950
 CC CC FT DOMAIN 951 1052
 CC CC FT DOMAIN 1118 1121
 CC CC FT DISULFD 74 129
 CC CC FT DISULFD 173 221
 CC CC FT DISULFD 270 320
 CC CC FT DISULFD 362 410
 CC CC FT CARBOHYD 73 73
 CC CC FT CARBOHYD 210 210
 CC CC FT CARBOHYD 326 326
 CC CC FT CARBOHYD 470 470
 CC CC FT CARBOHYD 489 489
 CC CC FT CARBOHYD 639 639
 CC CC FT CARBOHYD 715 715
 CC CC FT CARBOHYD 909 909
 CC CC FT CARBOHYD 1248 1300
 CC CC FT CONFLICT 168 168
 CC CC FT SEQUENCE 1461 AA; 159958 MW; 7AAEB9769635A21 CRC64;
 Query Match
 5.4%; Score 113; DB 1; Length 1461;

Best Local Similarity 23.1%; Pred. No. 0.37; Indels 72; Gaps 12;
Matches 65; Conservative 34; Mismatches 110;

18 NEVEEGPQNAVL-----KGSARPCNVSGMKLIMVLSDMVLSVREPIITN----- 69
Db 196 DRVTKLPGLMVLINATEGGDGLRCVYESGPP---KYDEVELKVLPPPEVLSLVLL 252
QY 70 -----DRFTSQRD--OGNFTSEMIIH 90
Db 253 KQSPPLVAVIGODVLPVCAAGLPTPIIKMKNEALDTSESLVLLAGS-----LEIS 308
QY 91 NVEPSDGNIRCSLQNSR--LHGSATLYOVYMGELFIPSNVLAENE-----POEYTCUP 144
Db 309 DVTEDDAGTYFCIADNGNETTEAQAELTVQAQPEFLKQPTNIYAHESMDIVFECEVTKRP 368
QY 145 SHWTLPDISW-ELGLVSHSSYFVPEPSPDQSAVSLIATLPQSNGLTCVATWKSILKA 203
Db 369 T-----PTVKVKNQMDVTPSDYFKIVEHNLQ-----VLGLVKSDEGFYCCIAENDVGN 419
QY 204 RKSATVNLTVIRCPODTGGGINIP-GVLSLPSLGSFLPIW 243
Db 420 QAGAQ-LILEHAPATGTGPLSPADVAVASIVSTFIKLTV 459
RESULT 11
ID OPCML HUMAN STANDARD: PRT: 345 AA.
AC 014982;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OBPCAM)
GN OPCML OR OBPCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Occipital cortex;
RX MEDLINE=95237612; PubMed=7721093;
RA Sherk K.B., Lee N.M.;
RT "Cloning, sequencing and localization to chromosome 11 of a cDNA
RT encoding a human opioid-binding cell adhesion molecule (OBPCAM).";
RL Gene 155:213-217(1995).
CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS. PROBABLY
CC INVOLVED IN CELL CONTACT.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: L34774; AAA36387.1;
CC Genew: HGNC:8143; OPCML.
DR MIM: 600632;
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_Like.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00410; IG_Like_1.
DR SMART: SM00408; IGC2_2.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.

FT SIGNAL 1 27
FT CHAIN 26 322
FT PROPEP 323 345
FT DOMAIN 50 122
FT DOMAIN 150 209
FT DOMAIN 237 303
FT DISULFD 57 115
FT DISULFD 157 202
FT DISULFD 244 296
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 140 140
FT CARBOHYD 285 285
FT CARBOHYD 293 293
FT CARBOHYD 306 306
FT CARBOHYD 322 322
SQ SEQUENCE 345 AA; 38007 MW; E7AD17B8A1A3FE4 CRC64;
Query Match 5.4%; Score 112.5; DB 1; Length 345;
Best Local Similarity 24.9%; Pred. No. 0.07; Indels 69; Gaps 16;
Matches 65; Conservative 32; Mismatches 95;

QY 26 NARVYKGSQARPCNVSGMKLIMVLSDMVLSVREPIITNDRFTSQRD 78
Db 44 NTVYRQGESATLRCITIDRVTRVAM-LNKSITLGNADKMSIDRVIIIVN---TPQY- 98
QY 79 OGNFTSEMIIHNEVEPSDGNIRCSLQ-----NSRLHGSATLYOVYMGELFIPSNVLA 132
Db 99 -----SIMIONVDYDGDGPTGCVQTDNHPKTSRVH---LIVQVPOQIMNISDITV 147
QY 133 AENEPEVTCV-----PSHWTLPDISWELGLVSHSSYF-----FVPEPSDQSAVSLIAT 184
Db 148 NESSSVTLICLAIGRPE-----PTVTVR-----FLSVKEGGGFVSEDEYLE-----ISDI 192
QY 185 TPQSNGLTCVATWKSILKARKSATVNLVY-----IRCPQDTGGGINIPGVY-----SSLPS 235
Db 193 KPDQSGTEECNAL-NVYAPDVRAKVIITVNPYPSKAKNTSVGQGIISCENSAVPM 251
QY 236 LGFSLPTWCK-----VGLGLAG 252
Db 252 AEFQ---WFKETRIATGLDG 269
RESULT 12
ID RAGE BOVIN STANDARD: PRT: 416 AA.
AC 028173;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Advanced glycosylation end product-specific receptor precursor
DE (Receptor for advanced glycosylation end products).
GN AGER OR RAGE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RF SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=92340547; PubMed=1378843;
RA Neepser M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
RA Elliston K., Stern D., Shaw A.;
RT "Cloning and expression of a cell surface receptor for advanced
RT glycosylation end products of proteins.";
RL J. Biol. Chem. 267:14998-15004(1992).
CC -1- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
CC RATE IN DIABETES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

Query Match	5.3%;	Score 111.5;	DB 1;	Length 345;
Best Local Similarity	25.0%;	Pred. No. 0.084;		
Matches .64;	Conservative 31;	Mismatches 96;	Indels 65;	Gaps 15

```

Db 99 -----SIMIONDVYDEGPTCSVQTDHNPKTSRVH---LIVQVPPQIMNISDVTV 147
Oy 133 AENEPEVTCL-----PSHWTLDPISWELGLVSHSSVY---VPPSPDSQASVIAL 184
Db 148 NEGSSVTLCLALGRPE-----PTVYWR-----HLSYKEQGVSEDEVYLE-----ISDI 192
Oy 185 TPSONGLTLCVATWKSLSKRSATVNLTV-----IRCPQDTGGGINIPGVL-----SSLPS 235
Db 193 KRDSGEYECASL-NDVAPAPVRKVKITVNPYISKAKNGVSGQGLISCASAVPM 251
Oy 236 LGFSLPTWGXVGLGLA 251
Db 252 AEFQ---WFKEDTRLA 264

RESULT 14
NEOL_CHICK STANDARD: PRT: 1443 AA.
AC 090610:
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neogenin (Fragment):
OS Gallus gallus (Chicken)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Embryonic brain;
RX MEDLINE=95105243; PubMed=7806578;
RA Vielmetter J., Roman J.M., Dreyer W.J.;
RT "Neogenin, an avian cell surface protein expressed during terminal
RT neuronal differentiation, is closely related to the human tumor
RT suppressor molecule deleted in colorectal cancer.";
RL J. Cell Biol. 127:2009-2020(1994).
CC -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DEVELOPMENTAL STAGE: IN RETINA, EXPRESSED ON GANGLION CELL FIBERS
CC AS SOON AS THEY BEGIN TO EXTEND THEIR AXONS.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07644; AAC59662.1; -
DR HSSP: P11276; 2MFN
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_repeat.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_Like.
DR Pfam: PF00041; fn3; 6.
DR Pfam: PF00047; ig; 4.
DR PRINTS: PRO0014; ENTPEPIT.
DR SMART: SM00060; FN3; 6.
DR SMART: SM00410; IG_Like; 2.
DR SMART: SM00408; IGC2; 2.
KW Transmembrane; Immunoglobulin domain; Glycoprotein; Repeat.
FT NON_TER 1

```

```

FT DOMAIN <1 1090 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1091 1111 POTENTIAL.
FT DOMAIN 1112 1443 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1112 1443 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 1112 1443 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 1112 1443 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 1112 1443 IG-LIKE C2-TYPE DOMAIN 4.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 1.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 2.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 3.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 4.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 5.
FT DOMAIN 1112 1443 FIBRONECTIN TYPE-III 6.
FT DISULFID 139 187 BY SIMILARITY.
FT DISULFID 236 286 BY SIMILARITY.
FT DISULFID 328 376 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1443 AA; 158050 MW; 558C679579C0E26 CRC64;

Query Match 5.3%; Score 111.5; DB 1; Length 1443;
Best Local Similarity 25.7%; Pred. No. 0.49; Mismatches 11;
Matches 48; Conservative 28; Indels 37; Gaps 11;

Oy 21 IEQPNARVILKGSQARFNCVSGW--KLIMWLSDMVLSVRMEIITNDFTSQR--77
Db 218 VROSSLTKVYGVNAVFPCC-VAGGFPPYVYVWVNG-----EEILIED--SERFA 264
Oy 78 -DQGNFTSEKTIINVEPSDGNRC--SLONSRLHGSAYLVQVMEELFIPSYNLVAE 134
Db 265 LRAGS-----LLISDVVEEDVGYTCTADNENETIEMQELAVYPPFLKRPANIYAE 320
Oy 135 NE---PCEVTLPSHWTLDPISM-ELGLLVSHSSVYFPEPSDQASVIALTPQSN 189
Db 321 SMDIVFEVEYVGRPT-----PTVWKVKNQGVVIPSDFYKIVEKHND---VLGLVKSDE 371
Oy 190 GTLTGVA 196
Db 372 GFYQCIA 378

RESULT 15
SMP_COTURN STANDARD: PRT: 620 AA.
ID SMP_COTURN
AC 092154:
DR 15-JUL-1998 (Rel. 36, Created)
DR 15-JUL-1998 (Rel. 36, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Schwann cell myelin protein precursor.
GN SMP.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-28; 120-132; 135-157 AND
RP 563-571.
RX MEDLINE=92153423; PubMed=1739462;
RA Dulac C., Tropak M.B., Cameron-Curry P., Rossier J., Marshak D.R.,
RA Roder J., le Douarin N.M.;
RT "Molecular characterization of the Schwann cell myelin protein, SMP:
RT structural similarities within the immunoglobulin superfamily.";
RL Neuron 8:323-334(1992).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED BY MYELINATING AND

```

Search completed: April 28, 2003, 18:09:41
Job time : 14.3692 secs

```
CC NONMYELINATING SCHWANN CELLS AND OLIGODENDROCYTES.
CC -1- DEVELOPMENTAL STAGE: FIRST SYNTHESIZED AT EMBRYONIC DAY 5, IT
CC REMAINS EXPRESSED BY CULTURED SCHWANN CELLS.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S83711; AAB21466.1; -.
CC DR HSSP: P56276; ITLK.
CC DR InterPro: IPR003006; Ig_MHC.
CC DR InterPro: IPR003598; Ig_C2.
CC DR InterPro: IPR003600; Ig_Like.
CC DR Pfam: PF00047; Ig_2.
CC DR SMART: SM00410; Ig_Like; 1.
CC DR SMART: SM00408; IGC2; 2.
CC KW Myelin; Glycoprotein; Cell adhesion; Transmembrane; Signal; Repeat;
CC KW Immunoglobulin domain.
CC FT SIGNAL 1 17
CC FT CHAIN 1 620
CC FT DOMAIN 18 516 SCHWANN CELL MYELIN PROTEIN.
CC FT TRANSMEM 517 536 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 537 620 POTENTIAL.
CC FT DOMAIN 537 620 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 28 106 IG-LIKE V-TYPE DOMAIN.
CC FT DOMAIN 151 223 IG-LIKE V-TYPE DOMAIN.
CC FT DOMAIN 253 311 IG-LIKE C2-TYPE DOMAIN 1.
CC FT DOMAIN 339 398 IG-LIKE C2-TYPE DOMAIN 2.
CC FT DOMAIN 424 495 IG-LIKE C2-TYPE DOMAIN 3.
CC FT DISULFID 35 164 IG-LIKE C2-TYPE DOMAIN 4.
CC FT DISULFID 40 99 BY SIMILARITY.
CC FT DISULFID 158 216 BY SIMILARITY.
CC FT DISULFID 260 304 BY SIMILARITY.
CC FT DISULFID 346 391 BY SIMILARITY.
CC FT DISULFID 420 429 BY SIMILARITY.
CC FT DISULFID 431 488 BY SIMILARITY.
CC FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 620 AA; 66943 MW; 004B3ECTEDC18FBA CRC64;

Query Match 5.38; Score 111; DB 1; Length 620;
Best Local Similarity 21.58; Pred. No. 0.19;
Matches 54; Conservative 33; Mismatches 84; Indels 80; Gaps 12;

QY 23 GPQNAVRLKGSQARFNCTV-SQGWKLIMWALSDWVLSVSRPMPEILITNDRFTSQRYDGG 81
DB 246 GP--TEVEGSDVELGCEAEGRPAFLISWFRGSEVL-----REPPGR 285
QY 82 NPTSEMIITHNVEPSDSNINRCSLONRSLHGSAYITVQVMELETPSYN--LVVAENDEPCE 139
DB 286 NL--RLLSNVGPDGGSFSCVAENRHRNRSIQLRLVAYAPRAPVINGSLWVVGDPVS 343
QY 140 VTCLPSHMTWLPDISWELGLVSHSSYFVPEPSDLOSAY---SIALT-----185
DB 344 VTCRAE-----SEPAALITVLRGKVMAMAAIYEDHVTMEMR 379
QY 186 ---PQNGTLTCAVATWMSLARKSATVNLTV-----TRCPDPTGGGINIPGVLSL 233
DB 380 PAREPDGTYSCVA--ENQHCASSTFNISVEYRPLVLPASRCTAG-GDSVRCVCMVNSI 436
QY 234 P--SLGFSLEPT 242
DB 437 PDSSLVELEPT 447
```

GenCore version 5.1.4-PS-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 18:09:54 ; Search time 19.0659 Seconds
(without alignments)
1946.300 Million cell updates/sec

Title: US-09-729-264-4

Perfect score: 2088
Sequence: 1 MVAGMERNRDPGSGSGNEV.....HPQASFLASPEKYSNTTV 386

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145.5	7.0	404	1	advanced glycosyla
2	136	6.5	1447	2	tumor suppressor p
3	134.5	6.4	1427	2	tumor suppressor -
4	133.5	6.4	871	1	protein-tyrosine k
5	129.5	6.2	148697	1	protein-tyrosine k
6	123.5	5.9	164	2	hypothetical prote
7	121	5.8	188	2	hypothetical prote
8	119.5	5.7	344	2	neurotrophin - rat
9	119.5	5.7	6642	2	protein-tyrosine k
10	119.5	5.7	1070	2	for-related recept
11	115.5	5.5	946	1	hypothetical prote
12	114.5	5.5	152	2	titin - mouse (fira
13	114.5	5.5	531	2	differentiation an
14	114.5	5.5	862	2	titin, cardiac mus
15	114.5	5.5	26926	1	opioid-binding cel
16	112.5	5.4	345	1	opioid-binding gly
17	112.5	5.4	416	1	CD22 homolog/B lym
18	112	5.4	868	2	advanced glycosyla
19	111.5	5.3	345	2	opioid-binding pro
20	111.5	5.3	1443	2	neogennin - chick
21	111.5	5.3	6805	2	titin - rabbit (fir
22	111.5	5.3	620	2	Schwann cell myeli
23	111	5.3	693	2	sodium-dependent p
24	110.5	5.3	338	2	opioid-binding pro
25	110.5	5.3	345	2	opioid-binding pro
26	110.5	5.3	4391	2	perlecan precursor
27	109.5	5.2	364	2	meloid cell surfa
28	109	5.2	3375	2	hypothetical prote
29	107	5.1	391	2	butyrophilin homol

30	106.5	5.1	487	2	butyrophilin - mou
31	106.5	5.1	802	2	mitogen and stress
32	106.5	5.1	841	2	killer cell inhibi
33	106.5	5.1	1272	2	neurofascin - chic
34	106	5.1	1092	2	neural cell adhesi
35	105.5	5.1	662	2	hypothetical prote
36	104.5	5.0	1177	2	hypothetical prote
37	104	5.0	764	2	irregular chlam C
38	104	5.0	2295	2	protein unc-52 (lm
39	104	5.0	5825	2	polypeptide - fava
40	103.5	5.0	2491	1	insulin-like growt
41	103.5	5.0	3707	2	heparan sulfate pr
42	102.5	4.9	890	1	protein-tyrosine k
43	102	4.9	721	2	hypothetical prote
44	101.5	4.9	215	2	hypothetical prote
45	101.5	4.9	1449	2	E2 glycoprotein pr

ALIGNMENTS

RESULT 1

161596
advanced glycosylation end-products receptor precursor - human
N:Alternate names: advanced glycosylation end product-binding protein, 35K, glycoprot
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C:Accession: 161596; B42879; S27968
R:Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko, G.
Genomics 23, 408-419, 1994
A:Title: Three genes in the human MHC class III region near the junction with the cla
nterpart of mouse mammary tumor gene int-3
A:Reference number: A55562; M01D:95137867; PMID:7835890
A:Accession: 161596
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-404 <RES>
A:Cross-references: GB:128769; NID:9561657; PIDN:BA05958.1; PID:9561659
R:Nepper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.;
J. Biol. Chem. 267, 14998-15004, 1992
A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation
A:Reference number: A42879; M01D:92340547; PMID:1378843
A:Accession: B42879
A:Molecule type: mRNA
A:Residues: 'G', '2-99', 'R', '101-404' <NEE>
A:Cross-references: EMBL:M91211; NID:9190845; PIDN:AAA03574.1; PID:9190846
A:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBIP:109438)
C:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly
cellular function, thus contributing to tissue lesions in diabetes.
C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C:Genetics:
A:Gene: GDB:AGER
A:Cross-references: GDB:306354; OMIM:600214
A:Map position: 6p21.3-6p21.3
A:Intons: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2
C:Function:
A:Description: neuronal receptor for amphoterin, a DNA-binding protein involved in ne
C:Superfamily: Advanced glycosylation end products receptor; immunoglobulin homology
C:Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-404/Product: advanced glycosylation end products receptor #status predicted <MAT
F:23-44/Domain: extracellular #status predicted <EXT>
F:31-101/Domain: immunoglobulin homology <IM1>
F:137-210/Domain: immunoglobulin homology <IM2>
F:252-303/Domain: immunoglobulin homology <IM3>
F:345-362/Domain: transmembrane #status predicted <TM>
F:363-404/Domain: intracellular #status predicted <INT>
F:25/81/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:38-99/144-208/259-301/Disulfide bonds: #status predicted

Query Match

7.0%; Score 145.5; DB 1; Length 404;

A: Molecule type: DNA
 A: Residues: 1-456, 'A', 466-871 <GAN2>
 A: Cross-references: EMBL:X86445; NID:9929723
 A: Experimental source: splice form 4
 C: Comment: For alternate splice forms see PIR:I48697.
 C: Genetics:
 A: Gene: nsk2
 A: Superfamily: MGI:103308
 C: Superfamily: mouse for-related receptor; immunoglobulin homology; protein kinase homology
 C: Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; tyrosine
 F: 1-21/Domain: signal sequence #status predicted <SIG>
 F: 22-871/Product: protein-tyrosine kinase nsk2, splice form 2 #status predicted <MAT2>
 F: 22-456, 'A', 466-871/Product: protein-tyrosine kinase nsk2, splice form 4 #status predicted
 F: 42-101/Domain: immunoglobulin homology <IMM1>
 F: 135-192/Domain: immunoglobulin homology <IMM2>
 F: 226-284/Domain: immunoglobulin homology <IMM3>
 F: 498-518/Domain: transmembrane #status predicted <TRM>
 F: 575-865/Domain: protein kinase homology <KIN>
 F: 583-591/Region: protein kinase ATP-binding motif
 F: 222,462/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 133.5; DB 1; Length 871;
 Best Local Similarity 21.6%; Pred. No. 0.045;
 Matches 73; Conservative 45; Mismatches 127; Indels 93; Gaps 17;

QY 4 G A M E N R D P P G S G N E Y I E G P O N A R V L K G S Q A R F N C T V S Q G M K - L I M W A L S D M V L S V R P 62
 D B 113 G A L O V K M K P - - - - - K I T R P P I N V K I I E G L K A V L P C T T M G N P K P S V S W I M K D - - - - - 158
 QY 63 M E P I T N D R F T S Q R Y D O G N F T S E M I T H N V E P S D S G I N R S L O N S R L H G S A Y - - - - - L T V 117
 D B 159 - N A L R E N S R I A L E - - - - - S G S L R I H N V O K E D A G Q Y C V A K N S L - - - G T A V S K L V K L E V 208
 QY 118 Q V M G E L F I P S N I L V A N E N P C E V T C - - - - - L P S H W T W L P D I S W - E L G L V S H S Y F V P E P 172
 D B 209 E V L G R I L R A P E S H N V T F G S F V T L R C T E I G P - - - - - V P T I S W I E N G A V S S G I O E S V K D 263
 QY 173 S D L O S A V S I L A L P O S N G T L T C V A T - - - - - W K S L K A R K S A T V N L T V I R - - - - - C P O D 219
 D B 264 R V I D S R L Q L F I T R P - - - - - G L Y T C I A T N K H G E K F S T A K A A T V S I A E W S K O K D S G Y C A O Y 320
 QY 220 T G G G I N T I G - - - - - V L S S L P S I G F S L P - - - - - T W G V G L G I A G T M L T P T C T I L I 264
 D B 321 R G E G V L M Q G P G E K M L V L P T T S H R D P E D A Q E L L I H T A M N E L - - - - - K A V S P L C R P A A 373
 QY 265 R C C C C R R C C G C N - - - - - C C C R C C - - - - - F C C R 287
 D B 374 E A L L C Y H L F L E C S P G V V P T P M P I C R E Y C L A V K E L F C A K 411

RESULT 5

I48697
 protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 1 - mouse
 N: Alternate names: receptor-type tyrosine kinase
 C: Species: Mus musculus (house mouse)
 C: Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
 C: Accession: I48697; S60740
 R: Ganju, P.; Wallis, E.; Brennan, J.; Keith, A. D.
 Oncogene 11, 281-290, 1995
 A: Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase
 A: Reference number: I48696; MUID:95349951; PMID:7624144
 A: Accession: I48697
 A: Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-881 <GAN1>
 A: Cross-references: EMBL:X86445; NID:9929725; PTDN:CA60166.1; PTD:9929726
 A: Experimental source: splice form 1
 A: Accession: S60740
 A: Molecule type: DNA
 A: Residues: 1-456, 'A', 466-881 <GAN2>
 A: Cross-references: EMBL:X86445; NID:9929725
 A: Experimental source: splice form 3

C: Comment: For alternate splice forms see PIR:I48696.

C: Genetics:
 A: Gene: nsk2
 A: Superfamily: MGI:103308
 C: Superfamily: mouse for-related receptor; immunoglobulin homology; protein kinase homology
 C: Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; tyrosine
 F: 1-21/Domain: signal sequence #status predicted <SIG>
 F: 22-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <MAT1>
 F: 22-456, 'A', 466-881/Product: protein-tyrosine kinase nsk2, splice form 3 #status predicted
 F: 42-101/Domain: immunoglobulin homology <IMM1>
 F: 135-192/Domain: immunoglobulin homology <IMM2>
 F: 226-284/Domain: immunoglobulin homology <IMM3>
 F: 498-518/Domain: transmembrane #status predicted <TRM>
 F: 575-865/Domain: protein kinase homology <KIN>
 F: 583-591/Region: protein kinase ATP-binding motif
 F: 222,462/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.4%; Score 133.5; DB 1; Length 881;
 Best Local Similarity 21.6%; Pred. No. 0.045;
 Matches 73; Conservative 45; Mismatches 127; Indels 93; Gaps 17;

QY 4 G A M E N R D P P G S G N E Y I E G P O N A R V L K G S Q A R F N C T V S Q G M K - L I M W A L S D M V L S V R P 62
 D B 113 G A L O V K M K P - - - - - K I T R P P I N V K I I E G L K A V L P C T T M G N P K P S V S W I M K D - - - - - 158
 QY 63 M E P I T N D R F T S Q R Y D O G N F T S E M I T H N V E P S D S G I N R S L O N S R L H G S A Y - - - - - L T V 117
 D B 159 - N A L R E N S R I A L E - - - - - S G S L R I H N V O K E D A G Q Y C V A K N S L - - - G T A V S K L V K L E V 208
 QY 118 Q V M G E L F I P S N I L V A N E N P C E V T C - - - - - L P S H W T W L P D I S W - E L G L V S H S Y F V P E P 172
 D B 209 E V L G R I L R A P E S H N V T F G S F V T L R C T E I G P - - - - - V P T I S W I E N G A V S S G I O E S V K D 263
 QY 173 S D L O S A V S I L A L P O S N G T L T C V A T - - - - - W K S L K A R K S A T V N L T V I R - - - - - C P O D 219
 D B 264 R V I D S R L Q L F I T R P - - - - - G L Y T C I A T N K H G E K F S T A K A A T V S I A E W S K O K D S G Y C A O Y 320
 QY 220 T G G G I N T I G - - - - - V L S S L P S I G F S L P - - - - - T W G V G L G I A G T M L T P T C T I L I 264
 D B 321 R G E G V L M Q G P G E K M L V L P T T S H R D P E D A Q E L L I H T A M N E L - - - - - K A V S P L C R P A A 373
 QY 265 R C C C C R R C C G C N - - - - - C C C R C C - - - - - F C C R 287
 D B 374 E A L L C Y H L F L E C S P G V V P T P M P I C R E Y C L A V K E L F C A K 411

RESULT 6

T24272
 hypothetical protein T01B7.8 - Caenorhabditis elegans
 C: Species: Caenorhabditis elegans
 C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C: Accession: T24272
 R: Sims, M.
 submitted to the EMBL Data Library, October 1995
 A: Reference number: T24272
 A: Accession: T24272
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-164 <MIL>
 A: Cross-references: EMBL:Z66499; PTDN:CA91301.1; GSPDB:GN00020; CESP:T01B7.8
 A: Experimental source: clone T01B7
 C: Genetics:
 A: Gene: CESP:T01B7.8
 A: Map position: 20/3; 90/2
 A: Introns: 20/3; 90/2

Query Match 6.2%; Score 129.5; DB 2; Length 164;
 Best Local Similarity 32.1%; Pred. No. 0.015;
 Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;

QY 157 L G L L V S S Y F P E P S D L O S A V S I L A L P O S N G T L T C V A T W K S L K A R K S A T V N L T V I R C 216
 D B 6 L A I L A I G T P I A V - - - - - S O V O S A V - - - - - L P V S T E L A T G T D V S T A S T A I D L I G N S S R V 57

OY 217 PDITGGINIPVLSLPSLGFSLPTWKGVLGLACTMILPTCTLTTRCCRRCCG 276
 Db 58 KRGGGCGCGCGC-----GCCGCGGGGG--CCCCRRCCCTC 101
 OY 277 --NCCC-RCFCRCR 287
 Db 102 CRTCCCTCCTCGR 115

RESULT 7

T15651

hypothetical protein C27A2.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15651

R:Nhan, M.

submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid C27A2.

A:Reference number: 218382

A:Accession: T15651

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-188 <NHA>

A:Cross-references: EMBL:U58760; NID:G1330384; PID:G1330389; PIDN:AAB00710.1; GSPDB:GNOC

A:Experimental source: strain Bristol N2; clone C27A2

C:Genetics:

A:Gene: CESP:C27A2.5

A:Map position: 2

A:Introns: 19/3; 91/2

Query Match

Best Local Similarity 5.9%; Score 122.5; DB 2; Length 188;
 Matches 19; Conservative 1; Mismatches 8; Indels 3; Gaps 2;

OY 260 CULTRCCGCCRRCCG--NCCC-RCFCRCR 287
 Db 86 CCCRRCCGCCRRCCCTCCTCCTCCTCGR 116

RESULT 8

156551

neurotrophin - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 19-May-2000

C:Accession: 156551

R:Struyk, A.F.; Canoll, P.D.; Wolfgang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.

J. Neurosci. 15, 2141-2156, 1995

A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed neur

A:Reference number: 156551; MUID:95198094; PMID:7891157

A:Accession: 156551

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-344 <RES>

A:Cross-references: EMBL:U06845; NID:9755184; PIDN:AAA67445.1; PID:9755185

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminal

Query Match

Best Local Similarity 5.8%; Score 121; DB 2; Length 344;
 Matches 59; Conservative 37; Mismatches 93; Indels 56; Gaps 13;

OY 26 NARLKSQARFNCYSQGLKLMALSDMNVLS-----VRMEPIITNDRFTSQRDQ 79
 Db 44 NVTYRQESATLCTIDNRTVRAVNRSTIYAGNDKWCIDPRVLLSN--TQTY-- 98
 OY 80 GGNFTSEMIHNVEPSDGNIRCSIQ-----NSRLHGAIVTVOMGELFIPSNLVYA 133
 Db 99 -----SIEIQNDVYDEGPYTCSTVDNHPKTSRYH-----LIVGSPKIVEISSDISIN 148
 OY 134 ENPECEVYTC---PSHWTLPDISWELGLVSHSSYFVPEPSDLOSASVIALTPQSN 189
 Db 149 EGNNSISYCIATGRPE-----PLVTWR---HISPAVGVSEDEYLE-----IQGITRQS 196

OY 190 GTLTCVATKSLKARSATVNLTVIRCP-----QDTGGGINIPVLSLPSLGFSL 240
 Db 197 GEYECSSAS--NDVAPVRRVNTVNNPYISAKGTGVPVGGKGLQCEANAVPSAERQ- 254
 OY 241 PTWKG 245
 Db 255 --WFK 257

RESULT 9

T29757

protein UNC-89 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999

C:Accession: T29757

R:Du, Z.; Lee, T.T.; Wilson, R.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C09D1.

A:Reference number: 220679

A:Accession: T29757

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6642 <DUZ>

A:Cross-references: EMBL:AF003131; PIDN:AAB54132.1; GSPDB:GN00019; CESP:unc-89

A:Experimental source: strain Bristol N2; clone C09D1

C:Genetics:

A:Gene: CESP:unc-89

A:Map position: 1

A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/3; 352/3; 426/2; 454/1; 500/1; 537/1

3/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match

Best Local Similarity 5.7%; Score 119.5; DB 2; Length 6642;
 Matches 51; Conservative 39; Mismatches 93; Indels 35; Gaps 8;

OY 20 VIEGQNAKRVKGSQARFNCYSQGL--KLIMALSDMVLVSRMEPIITNDRFTSQR 77
 Db 2077 VVDGKSVYTKETETAEFKATIS--GPPAPVTKWTINEKIYESRTITTKTEDVY---- 2131
 OY 78 DGGNFTSEMIHNVEPSDGNIRCSLONSRLHGSATVTVOMGELFIS-----VNLVY 132
 Db 2132 -----LKISNAKLEQIGYKAVTAQNSAGDSQADIKPPNVKARKFSQGLTDKVA 2182
 OY 133 AENPECEVYTC---PSHWTLPDISWEL--GLVSHSSYFVPEPSDLOSASVIALTPQ 188
 Db 2183 DEGPRLWNLDELDPSPGCT---EVSMLNGLPRLTKSDTVQVVDHGGYHVTIAAKREM 2239
 OY 189 NGTLTCVATKSLKARSATVNLTVIRCPD 219
 Db 2240 SGTLTAKAKNAAGCETSAKVTVNNGNKKRPEFVQADQN 2277

RESULT 10

JC4593

protein-tyrosine kinase-related receptor PTK7 precursor - human

N:Alternate names: receptor protein tyrosine kinase-like protein (RPTK)

C:Species: Homo sapiens (man)

C>Date: 16-Apr-1996 #sequence_revision 24-May-1996 #text_change 24-Sep-1999

C:Accession: JC4593

R:Park, S.K.; Lee, H.S.; Lee, S.T.

J. Biochem. 119, 235-239, 1996

A:Title: Characterization of the human full-length PTK7 cDNA encoding a receptor prot

A:Reference number: JC4593; MUID:97037064; PMID:8882711

A:Accession: JC4593

A:Molecule type: mRNA

A:Residues: 1-1070 <PAR>

A:Cross-references: GB:U40271; NID:G1322231; PIDN:AAC50484.1; PID:G1322232

C:Comment: This protein is a member of receptor protein tyrosine kinase family, but p

C:Genetics:

A:Gene: GDB:PTK7

A:Cross-references: GDB:134760; OMIM:601890

A:Map position: 6p21.1-6p12.2

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

Db 264 RYDLSRLQFLTKP---GLYTICIAIANKGEKSTAKAAATISIAEWSPKDKNGYCAOY 320
 QY 224 ----INIGVLSLPSLGSFSL-----PTWGVGLAGLMTLPTCTLTIRCC 267
 Db 321 RGEVCNAVIAKDALVFLMTSYADPEAOELLVHTAMNEL-----KVSPVCRPAEAL 373
 QY 268 CCRRCRCGCGN-----CCCRCC-----FCCRRKGRFRIOPKSEKTEKETETE 312
 Db 374 LCNHIFQECSPGVPLPIPICREYCLAVKEIFCAKE-----WLVMEKTHRGLYRSEMHLL 429
 QY 313 SGENSGVNSDEQKTTDASLPKSCESDPEQRNNSCGPHORADRP-----PRPASHP 368
 Db 430 SYECSKLPMSMHMDPTACARLP-----HLDYKNENLKTFFP--MTSSKPSVDIPLPSSS 482
 QY 369 QASEFNLA 375
 Db 483 SSSFSVS 489

RESULT 4

US-08-644-271-29
 ; Sequence 29, Application US/08644271
 ; Patent No. 5814478

GENERAL INFORMATION:

APPLICANT: Valenzuela, et al.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
 TITLE OF INVENTION: AND LIGANDS
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 STREET: 777 Old Saw Mill Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/644,271
 FILING DATE: 10-MAY-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USN 60/008,657
 FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Covert, Robert J
 REGISTRATION NUMBER: 36,108
 REFERENCE/DOCKET NUMBER: REG 195A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 914-345-7400
 TELEFAX: 914-345-7721

TELEX:

INFORMATION FOR SEQ ID NO: 29:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 869 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-644-271-29

Query Match

6.3%; Score 132.5; DB 2; Length 869;
 Best Local Similarity 20.8%; Pred. No. 0.0018;
 Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;

QY 4 GAMEDRPPGSGSGNEVEIGPONARVLKGSQARENCTVSQGWK-LIMMALSDMVYLSVRP 62
 Db 113 GALQVAKMP-----KITRPINVKIIEGLKAVLPCPTMGNPKRPSVMIKD----- 158
 QY 63 MEPIITNDFTSORYDQGNFTSEMIINHVEPSDSGNIRCSLQNSRLGSAV-LTVQVMG 121

Db 159 -SPLENSRIAVLE-----SGSLRIHNVOKEPACQRCVANKSL--GTAYSKVYKLEY 208
 QY 122 ELFTPSVNLVAENEP-----CEVTCLEPSHHTMLPDLSM-ELGLVSHSSYFVPEP 172
 Db 209 EVFARILRAPESHNTFGSFVTLHCTANGIP-----VPTIWIENGNAVSSGSIQESVKD 263
 QY 173 SDLSQASVSLALTPQSNGLTCVAT---WKSIAKRSATVNLTVIRCPDGTGGG----- 223
 Db 264 RYDLSRLQFLTKP---GLYTICIAIANKGEKSTAKAAATISIAEWSPKDKNGYCAOY 320
 QY 224 ----INIGVLSLPSLGSFSL-----PTWGVGLAGLMTLPTCTLTIRCC 267
 Db 321 RGEVCNAVIAKDALVFLMTSYADPEAOELLVHTAMNEL-----KVSPVCRPAEAL 373
 QY 268 CCRRCRCGCGN-----CCCRCC-----FCCRRKGRFRIOPKSEKTEKETETE 312
 Db 374 LCNHIFQECSPGVPLPIPICREYCLAVKEIFCAKE-----WLVMEKTHRGLYRSEMHLL 429
 QY 313 SGENSGVNSDEQKTTDASLPKSCESDPEQRNNSCGPHORADRP-----PRPASHP 368
 Db 430 SYECSKLPMSMHMDPTACARLP-----HLDYKNENLKTFFP--MTSSKPSVDIPLPSSS 482
 QY 369 QASEFNLA 375
 Db 483 SSSFSVS 489

RESULT 5

US-09-077-955-33
 ; Sequence 33, Application US/09077955A
 ; Patent No. 6413740

GENERAL INFORMATION:

APPLICANT: Valenzuela et al., David M.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
 FILE REFERENCE: REG195-B-PCT-US
 CURRENT APPLICATION NUMBER: US/09/077,955A
 CURRENT FILING DATE: 1998-09-10
 EARLIER APPLICATION NUMBER: PCT/US96/20696
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 08/644,271
 EARLIER FILING DATE: 1996-05-10
 EARLIER APPLICATION NUMBER: 60/008,657
 EARLIER FILING DATE: 1995-12-15
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 33
 LENGTH: 869
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-077-955-33

Query Match

6.3%; Score 132.5; DB 4; Length 869;
 Best Local Similarity 20.8%; Pred. No. 0.0018;
 Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;

QY 4 GAMEDRPPGSGSGNEVEIGPONARVLKGSQARENCTVSQGWK-LIMMALSDMVYLSVRP 62
 Db 113 GALQVAKMP-----KITRPINVKIIEGLKAVLPCPTMGNPKRPSVMIKD----- 158
 QY 63 MEPIITNDFTSORYDQGNFTSEMIINHVEPSDSGNIRCSLQNSRLGSAV-LTVQVMG 121
 Db 159 -SPLENSRIAVLE-----SGSLRIHNVOKEPACQRCVANKSL--GTAYSKVYKLEY 208
 QY 122 ELFTPSVNLVAENEP-----CEVTCLEPSHHTMLPDLSM-ELGLVSHSSYFVPEP 172
 Db 209 EVFARILRAPESHNTFGSFVTLHCTANGIP-----VPTIWIENGNAVSSGSIQESVKD 263
 QY 173 SDLSQASVSLALTPQSNGLTCVAT---WKSIAKRSATVNLTVIRCPDGTGGG----- 223
 Db 264 RYDLSRLQFLTKP---GLYTICIAIANKGEKSTAKAAATISIAEWSPKDKNGYCAOY 320
 QY 224 ----INIGVLSLPSLGSFSL-----PTWGVGLAGLMTLPTCTLTIRCC 267

Db 321 RGEVNAVLAADLVFINTSADPEEAQELLVHTANNEI-----KVSPICORAAEAL 373
 QY 268 CRRRCGCGN-----CCRCRC-----FCGRKRGRIOPQKSKSEKTNKETE 312
 Db 374 LCNHIFQECSPGVYPTPIPICREYCLAVKELFCAKE---WLMERKTRGLRSEMLL 429
 QY 313 SGNNSGYNSDEQKTTDTASLPPKSCSSDDEQRNNSCGPPHQRADQRP---PRPASHP 368
 Db 430 SVPECSXLPMSHMDPTACARLP-----HLDYKNENLKTRPP--MSSKRSVIDPMLPSS 482
 QY 369 QASFNLA 375
 Db 483 SSSRSVS 489

RESULT 6
 ; Sequence 3, Application US/08977767
 ; Patent No. 5972684
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Greenwald, Sara
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/977,767
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0423 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; TELEX:
 ; INFORMATION FOR SRO ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1345 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 1532042
 ; US-08-977-767-3

Query Match 6.2%; Score 128.5; DB 2; Length 1345;
 Best Local Similarity 33.0%; Pred. No. 0.0077;
 Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5;
 QY 190 GTTCATATKSLAKRSATVNLTVIRCPDGTGGI-----NIPGVLSLPSLGFSLPTWGK 245
 Db 414 GTCTGTGT-----GC-CGTGGAAAGCGTCAAGAGCCCGCGATGTGGA 455

QY 246 VGLAGLMTLT-PTCTLTIRCCCRRCGCCNCCRC-----CFCC 286
 Db 456 CGTGAAGAGGCTCTCTATGACCCCTTCCTGGCCCTCTGAGACATGACACC 507
 RESULT 7
 ; US-09-651-200-2
 ; Sequence 2, Application US/09651200
 ; Patent No. 6429303
 ; GENERAL INFORMATION:
 ; APPLICANT: Green et al.
 ; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 ; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 ; FILE REFERENCE: 15966-562 (CURA-62)
 ; CURRENT APPLICATION NUMBER: US/09/651,200
 ; PRIOR APPLICATION NUMBER: 2000-08-30
 ; PRIOR APPLICATION NUMBER: 1999-09-03
 ; PRIOR APPLICATION NUMBER: 60/152383
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/172909
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/183578
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 340
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-651-200-2

Query Match 5.9%; Score 123.5; DB 4; Length 340;
 Best Local Similarity 21.3%; Pred. No. 0.0033;
 Matches 78; Conservative 50; Mismatches 130; Indels 109; Gaps 17;
 QY 11 PPGSGGNEVIEGPQNAHV-LKGSQARFNCTVS--QGKLT-----IMWALSDMYVLSVRPM 63
 Db 45 PQRPTGAVEVQVEDPVVALVGTDAHLHCSFSPREPFSLTQNLMLQTLDTQVLY----- 100
 QY 64 EPIINDRETSQRYDCCGN-----TSMITHNVPSDSGNIRCSLONSRL 109
 Db 101 -----HSFIEGR-DQGSATANRTALFPDLLAOGNMSLRQRYRVADSESFCEV-SIRD 152
 QY 110 HGSAYLVQVYVGEFLFIPSVNLV---VAENPEVCLPSHWMUP--DISMELG---L 160
 Db 153 FGSAAVSLQVAAPYSKSMLEPNKDLRPGDTVTIIC--SYRIGTPEAEVFWQDQGVPL 210
 QY 161 VSHSSYFVEPEPDLQSAVSLALTPQSNGLTQVATMKSLSAKRSATVNLTVIRCP--- 217
 Db 211 TGNVTTQMANEGLFVHSHVVLGVANGTYS-----LVANPVLYQ 252
 QY 218 QDTGGGINIGVLSLSLGFSLPTWKGVLGLAGTMTLTPTCTLTIRCCCRRCGCCN 277
 Db 253 QDAHGSVTITGQPMTPPEAL-----WYVGLSVCLIALLV----- 288
 QY 278 CCRRCGCCRRKRGFRIRIQFKSKSEKTNKETEESGNSGNSYNSDEQKTTDTASLPPKS 337
 Db 289 ---ALAFVLCWRK-----IKOSCEEENAGAEDDQ-----EEGSKTLQPLKH 328
 QY 338 CESSDPE 344
 Db 329 SDSKEDD 335

RESULT 8
 ; US-09-651-200-4
 ; Sequence 4, Application US/09651200
 ; Patent No. 6429303
 ; GENERAL INFORMATION:
 ; APPLICANT: Green et al.
 ; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 ; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 ; TITLE OF INVENTION: Polypeptides Encoded Thereby

FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 441
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-200-4

Query Match
Best Local Similarity 21.3%; Pred. No. 0.0047;
Matches 78; Conservative 50; Mismatches 130; Indels 109; Gaps 17;

5.9%; Score 123.5; DB 4; Length 441;
11 PPGSGNEVEIEGPONARY-LKSGARFNCVTS--QGKRL---IMALSDMYLSVRPM 63
146 PQRSPTGAVEVQVPEDEPVVALVGTDTALHCSFSPGFSIAQLNLIWQITDKQLV---- 201
64 EPIITNDRTSORYDOGNF-----TSEMIIHNEPDSGNIKCSLQNSRL 109
202 -----HSTFTEGR-DQGSAYANRTALFPDLAOGNNSLRLOVRVADESGFTCEV-SIRD 253
110 HGSAYLVQVMGELFIPSVNLV---VAENPECVTCPLPSHTWLP--DISMELGL---L 160
254 FGSAAVSLQVAAPYKPSMTLEPKDLRPDGYTITC--SSYRGYPBAEVFWDDGGVPL 311
161 VSHSSYFVEPEPSDLOSASVIALTPQSNCTLCVATWKSLSKRSATVNLVIRCP--- 217
312 TGNVTTISQMANEQGLFDVHSLRVLVGANGTYSCL-----LVKRPVILQ 353
218 QDTGGGINIRGVLSLSLPSLGFSLPTWKGVLGLAGTMLLPCTCLITRCCCRRCGCGN 277
354 QDAHGSVLTIGQPMTPPEAL---WVTGLSVCLLALLV-----EGESKTALQPLKH 389
278 CCCRCFCRRKRGFRIOFKKSEKTKETETESGNSGNSDQKTTDTASLPKS 337
390 ---ALAFVCWRK-----IKOSCEENAGAEIDQD-----EGESKTALQPLKH 429
338 CESSDPE 344
430 SDSKEDD 436

RESULT 9
US-09-651-200-6
Sequence 6, Application US/09651200
Patent No. 6429303
GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 534
TYPE: PRT

ORGANISM: Homo sapiens
US-09-651-200-6

Query Match
Best Local Similarity 21.3%; Pred. No. 0.0075;
Matches 78; Conservative 49; Mismatches 131; Indels 109; Gaps 17;

5.9%; Score 122.5; DB 4; Length 534;
11 PPGSGNEVEIEGPONARY-LKSGARFNCVTS--QGKRL---IMALSDMYLSVRPM 63
239 PQRSPTGAVEVQVPEDEPVVALVGTDTALHCSFSPGFSIAQLNLIWQITDKQLV---- 294
64 EPIITNDRTSORYDOGNF-----TSEMIIHNEPDSGNIKCSLQNSRL 109
295 -----HSTFTEGR-DQGSAYANRTALFPDLAOGNNSLRLOVRVADESGFTCEV-SIRD 346
110 HGSAYLVQVMGELFIPSVNLV---VAENPECVTCPLPSHTWLP--DISMELGL---L 160
347 FGSAAVSLQVAAPYKPSMTLEPKDLRPDGYTITC--SSYRGYPBAEVFWDDGGVPL 404
161 VSHSSYFVEPEPSDLOSASVIALTPQSNCTLCVATWKSLSKRSATVNLVIRCP--- 217
405 TGNVTTISQMANEQGLFDVHSLRVLVGANGTYSCL-----LVKRPVILQ 446
218 QDTGGGINIRGVLSLSLPSLGFSLPTWKGVLGLAGTMLLPCTCLITRCCCRRCGCGN 277
447 QDAHGSVLTIGQPMTPPEAL---WVTGLSVCLLALLV-----EGESKTALQPLKH 482
278 CCCRCFCRRKRGFRIOFKKSEKTKETETESGNSGNSDQKTTDTASLPKS 337
483 ---ALAFVCWRK-----IKOSCEENAGAEIDQD-----EGESKTALQPLKH 522
338 CESSDPE 344
523 SDSKEDD 529

RESULT 10
US-09-651-200-24
Sequence 24, Application US/09651200
Patent No. 6429303
GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Sequence
OTHER INFORMATION: mz5020.protein from Figure 4.
US-09-651-200-24

Query Match
Best Local Similarity 21.3%; Pred. No. 0.0075;
Matches 78; Conservative 49; Mismatches 131; Indels 109; Gaps 17;

5.9%; Score 122.5; DB 4; Length 534;
11 PPGSGNEVEIEGPONARY-LKSGARFNCVTS--QGKRL---IMALSDMYLSVRPM 63
239 PQRSPTGAVEVQVPEDEPVVALVGTDTALHCSFSPGFSIAQLNLIWQITDKQLV---- 294

```

QY 64 EPIITNDRTSQRYDOGNF-----TSEMTHNVEPSDSGNIRCSLQNSRL 109
Db 295 -----HSFTEGR-DOGSVAMNRTALPDLAQNASLQLQVRVADSGSFCEP-SIRD 346
QY 110 HGSAYLVQVWGLFIPSNLV-----VAENECEVTCPLSHMTWLP--DISMELGL--L 160
Db 347 FGSANVSLOYAARYSKPSMTLEPNKDLRPGDVITTC--SSRYGPEAEVEMODQGVPL 404
QY 161 VSHSSYFVPEPSDLOSASVIALTPQSNGLTCVATWKSILKARKSATVNLVIRCP--- 217
Db 405 TGVVTTQSOMANEGLFDVHVSILRVILGANGTYS------LVNRPVLO 446
QY 218 QDTGGGINIPVLSLPSLGSFLPTWKGVLGLAGTMLLTPTCLTTRCCRRCCGCGN 277
Db 447 QDHGSGVTTTGQPMTEPPPAL-----WVTVGLSVCLIALLV----- 482
QY 278 CCCRCFCRRKRGFRIOKSEKENTKETETESGNGNSDDEKTTDTASLPKPS 337
Db 483 ---ALAFVCMRK-----IKOSCEENAGAEODG-----HEGSKTALQPLKH 522
QY 338 CESSDPE 344
Db 523 SDSKEDD 529

```

RESULT 11

```

US-09-062-365-1
; Sequence 1, Application US/09062365
; Patent No. 6465422
; GENERAL INFORMATION:
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Stern, David
; TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
; FILE REFERENCE: 55424
; CURRENT APPLICATION NUMBER: US/09/062,365
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Human
US-09-062-365-1

```

```

Query Match
Best Local Similarity 5.8%; Score 121.5; DB 4; Length 332;
Matches 57; Conservative 27; Mismatches 69; Indels 69; Gaps 11;

```

```

QY 64 EPIITNDRTS-----ORYDOGNFT--SEMTHNVEPSDSGNIR-----CSLQNSRLHGS 112
Db 140 KPLVPNKGVSVKQTRRHPTGLFTLQSLM---VTPARGDPRTFSCSFSPGLPRHR 196
QY 113 AYLVOVWGLFIP-----SYNLVAENEP-----CEVTCPLSHMTWLPDIS 154
Db 197 ALRTAPLQPRVWEVPLEEVLV---EPGGAVAPGVTILCEVPAPPS-----PQIH 248
QY 155 WELGLVSHSSYFVPEPSDLOSASVIALTPQSNGLTCVATWKSILKARKSATVNLVIR 214
Db 249 WMKD-----CVPLPLPSPVLLPLPEIGPDQGTSCVATHSHGPESSRAVSISII 299
QY 215 RCPDGTGGGINIPVLSLPSLGSFLPTWKGVLGLAGTMLL 256
Db 300 E-PGEEG-----PTAGSVGSGSLGLTAL 321

```

RESULT 12

```

US-09-540-245A-15
; Sequence 15, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas

```

```

; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,245A
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/065,544
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-540-245A-15

```

```

Query Match
Best Local Similarity 5.7%; Score 120; DB 4; Length 1395;
Matches 66; Conservative 29; Mismatches 101; Indels 56; Gaps 15;

```

```

QY 24 PQNARVLKGSQARENCVTSQG--WKLIMW-----ALSDMVLSVRPEPIITNDRTSOR 76
Db 161 PKDTRVAKGETALLCEGPKPIPEPTLWIKDGVPLDLLAMSGASSRV-----R 211
QY 77 YDQGNFTSEMIHNEPSDSGNIRCSLQ--SRLHGSAYLVQVWGLFIPSNLVYA 133
Db 212 IVDGN-----LLISNVEPIDEGNKCAIONLVGRSSYAKLIVQKPFEMKPKQVWL 267
QY 134 ENEP-----CEVTCPLSHMTWLPDISW---ELGLVSHSSYFVPEPSDLOSASVIALTP 186
Db 268 YGQATFHCISVGDP-----PKVLKKEGNIIVSARILHD-----EKSLTISNTP 316
QY 187 QSNGLTCVA--TWKSLKARKSATV---NLTVIRCPDGTGGGINIPVLSLPSLGSFL 240
Db 317 TDEGTVCEAHNNVQGISARASILVHAPNFT--KRPSSKKVGLN--GVV-QLPCMASGN 371
QY 241 PT-----NGKXGL 248
Db 372 PPSVFWTKGCV 383

```

RESULT 13

```

US-08-374-834-1
; Sequence 1, Application US/08374834
; Patent No. 5656473
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/374,834
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/095,658
; FILING DATE: 21-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Covert, Robert J.
; REGISTRATION NUMBER: 36,108

```

REFERENCE/DOCKET NUMBER: REG 190A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 345-7400
 TELEFAX: (914) 345-7721
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 868 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-374-834-1

Query Match 5.7%; Score 118; DB 1; Length 868;
 Best Local Similarity 21.0%; Pred. No. 0.037;
 Matches 83; Conservative 42; Mismatches 151; Indels 120; Gaps 19;

QY 4 GAEHRDPGSGSGNEVEIEGPNARVLKSGOARFNCVSGWK-LIMALSDMYLVSRP 62
 DB 113 GALQVKKMP-----KTRPPINVKIIEGLKAVLPCTTMGNPKPSVMIKGDALRE--- 163
 QY 63 MEPIITNDPFTSQRYDQGNFTSEMIINHPSPDSGNIRCSLONSRLHGSAY-LTVQVWG 121
 DB 164 -----NSRIAVLE-----SGSLRIHNVQKEDAGQRCVAKNSL--GTAYSKLVKLEV 208
 QY 122 ELFIPTVNLVVAENP-----CEVTCPLPSHMTWLPDISW-ELGLVSHSSYFVEPEP 172
 DB 209 EVFARILRAPESHNVTFEGSFVTLRCAIGMP-----VPTISWINGNAVSSGSIQENVKD 263
 QY 173 SDIOSAVSIILALTPQSNGLTCVAT---WKSILKARKSATVNLTVIR-----CPQD 219
 DB 264 RVIDSRLOFLITKP---GLYTCIATNKHGKEFTAKAATVSIAMWSKQSKESKGYCAQY 320
 QY 220 TGGGINIPGVLSLPSLGSFLP-----TWKRVGLAGLMTLPTCTLLIRRC 267
 DB 321 RGEVCAVLYKDSLVEFNTSYDPDEAQLIHITAMNEL-----KAVSPLCRPAEAL 373
 QY 268 CCRRCGCCGN-----CCCRCC-----FCCRRKRGFRIOFOKSEKETE 312
 DB 374 LCNHLFQECSPGVLPMPICREYCLAVKELFCA-----KEMLMGKTH 418
 QY 313 SGNSGYNSDEQKTTDTASLPPKSCS-----SDP 343
 DB 419 RGLYRSGMH-----FLPVPECSKLPSMHQDP 444

RESULT 14
 US-08-644-271-1
 Sequence 1, Application US/08644271
 Patent No. 5814478
 GENERAL INFORMATION:
 APPLICANT: Valenzuela, et al.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 STREET: 777 Old Saw Mill Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/644,271
 FILING DATE: 10-MAY-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 60/008,657

FILING DATE: 15-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Cobert, Robert J
 REGISTRATION NUMBER: 36,108
 REFERENCE/DOCKET NUMBER: REG 195A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 914-345-7400
 TELEFAX: 914-345-7721
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 868 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-644-271-1

Query Match 5.7%; Score 118; DB 2; Length 868;
 Best Local Similarity 21.0%; Pred. No. 0.037;
 Matches 83; Conservative 42; Mismatches 151; Indels 120; Gaps 19;

QY 4 GAEHRDPGSGSGNEVEIEGPNARVLKSGOARFNCVSGWK-LIMALSDMYLVSRP 62
 DB 113 GALQVKKMP-----KTRPPINVKIIEGLKAVLPCTTMGNPKPSVMIKGDALRE--- 163
 QY 63 MEPIITNDPFTSQRYDQGNFTSEMIINHPSPDSGNIRCSLONSRLHGSAY-LTVQVWG 121
 DB 164 -----NSRIAVLE-----SGSLRIHNVQKEDAGQRCVAKNSL--GTAYSKLVKLEV 208
 QY 122 ELFIPTVNLVVAENP-----CEVTCPLPSHMTWLPDISW-ELGLVSHSSYFVEPEP 172
 DB 209 EVFARILRAPESHNVTFEGSFVTLRCAIGMP-----VPTISWINGNAVSSGSIQENVKD 263
 QY 173 SDIOSAVSIILALTPQSNGLTCVAT---WKSILKARKSATVNLTVIR-----CPQD 219
 DB 264 RVIDSRLOFLITKP---GLYTCIATNKHGKEFTAKAATVSIAMWSKQSKESKGYCAQY 320
 QY 220 TGGGINIPGVLSLPSLGSFLP-----TWKRVGLAGLMTLPTCTLLIRRC 267
 DB 321 RGEVCAVLYKDSLVEFNTSYDPDEAQLIHITAMNEL-----KAVSPLCRPAEAL 373
 QY 268 CCRRCGCCGN-----CCCRCC-----FCCRRKRGFRIOFOKSEKETE 312
 DB 374 LCNHLFQECSPGVLPMPICREYCLAVKELFCA-----KEMLMGKTH 418
 QY 313 SGNSGYNSDEQKTTDTASLPPKSCS-----SDP 343
 DB 419 RGLYRSGMH-----FLPVPECSKLPSMHQDP 444

RESULT 15
 US-09-077-955-1
 Sequence 1, Application US/09077955A
 Patent No. 6413740
 GENERAL INFORMATION:
 APPLICANT: Valenzuela et al., David M.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
 FILE REFERENCE: REG195-B-PCT-US
 CURRENT APPLICATION NUMBER: US/09/077,955A
 CURRENT FILING DATE: 1998-09-10
 EARLIER APPLICATION NUMBER: PCT/US96/20696
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 08/644,271
 EARLIER FILING DATE: 1996-05-10
 EARLIER APPLICATION NUMBER: 60/008,657
 EARLIER FILING DATE: 1995-12-15
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 868
 TYPE: PRT
 ORGANISM: Rattus sp.

US-09-077-955-1

Query Match 5.7%; Score 118; DB 4; Length 868;

Best Local Similarity 21.0%; Pred. No. 0.037;

Matches 83; Conservative 42; Mismatches 151; Indels 120; Gaps 19;

```

QY 4 GAMENDPPGSGGNEVIGPONARVYKGSQARFNCTVSQGWK-LIMWALSDMVVLSYRP 62
DB 113 GALQVMMKP-----KITRPPINVKIIEGLKAVLPCTTMGNPKPSVSWIKGDSALRE--- 163
QY 63 MEPIITNDPFTSQRYDQGNFTSEMIINHVEPDSGNIRCSLQNSRLHGSAY-LTVQVMG 121
DB 164 -----NSRIAYLE-----SGSLRHHVQKEDAGQYRCVAKNSL--GTAYSKLVKLEV 208
QY 122 ELFIPTSVNLVVAENP-----CEVTCLPSSHMTWLPDISW-ELGLVSHSSYYFVPEP 172
DB 209 EYFARILRAPESHNVTFGSGFVTLRCTAIGMP-----VPTISMIENGNAVSSGSIQENVKD 263
QY 173 SDIOSAVSILALPQSNCTLCQVAT---WKSIRAKRSATVNLVIR-----CPQD 219
DB 264 RYIDSRLQLFITRP---GLYTCTIATNKHGKEFTAKAAATVSIAWSKSQKESKGYCAQY 320
QY 220 TGGGINIPGVLSLPSLGSFLP-----TWGKVGGLAGTMLLPPTCTLTIRCC 267
DB 321 RGEVCDAYLVKDSLVEFNTSYDPDEEAQELIHITAMNEL-----KAVSPICRPAAEAL 373
QY 268 CCRRCRCGCGN-----CCCRCC-----FCCRKRKGFRIOPQKSEKERTNKETETE 312
DB 374 LCNHLPQECSPGVLPPTPMPICREYCLAVKELFCA-----KEMWLAMEGKTH 418
QY 313 SGNENSGYNSDEOKTTDTASLPKSCES-----SDP 343
DB 419 RGLYRSGMH-----FLPVPECSKLPMSHODP 444

```

Search completed: April 28, 2003, 21:12:19
 Job time : 19.0485 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:24:31 ; Search time 38.4662 Seconds
(without alignments)
1337.141 Million cell updates/sec

Title: us-09-729-264-4

Perfect score: 2088
Sequence: 1 MAGANENDPPGSGSGNFV.....HPOASFNLASPEKSVNTTVV 386

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2088	100.0	386	23	AAU75541 Human B7-1-like prot
2	2004	96.0	382	23	AAU75540 Human B7-1-like prot
3	2000	95.8	386	23	AAU75542 Human B7-1-like prot
4	1576	75.5	377	23	AAU75543 Human B7-1-like prot
5	1318	63.1	463	22	ABG28169 Novel human diago
6	913.5	43.8	370	23	AAU75544 Mouse B7-1-like prot
7	572.5	27.4	631	23	AAU75547 Rat B7-1-like protei
8	566.5	27.1	270	23	AAU75545 Mouse B7-1-like prot
9	485	23.2	223	23	AAU75546 Mouse B7-1-like prot
10	145.5	7.0	404	22	AA81925 Extracellular c1r

11	145.5	7.0	404	23	AAE23219 Human receptor for
12	145.5	7.0	404	23	AAU77543 Human PAGE protein
13	145.5	7.0	404	23	AA848745 Human PAGE protein
14	138	6.6	1496	20	AA81030 Melanoma associate
15	138	6.6	1496	21	AAV70467 Human p53 target m
16	138	6.6	1496	22	AB811587 Human peroxidasin
17	136	6.5	1447	16	AA868553 Deleted in colorec
18	136	6.5	1447	16	AAV33498 Human DCC protein.
19	136	6.5	1447	22	AA833498 Deleted in colorec
20	136	6.5	1447	22	AA850693 Human UNC-40 prote
21	136	6.5	1447	22	AA813144 Deleted in colorec
22	133.5	6.4	475	19	AA862575 Alternately splic
23	133.5	6.4	475	19	AA894982 Alternately splic
24	133.5	6.4	863	19	AA862569 Mouse receptor tyr
25	133.5	6.4	863	19	AA862583 Mouse receptor tyr
26	133.5	6.4	863	19	AA862583 Mouse receptor tyr
27	133.5	6.4	863	19	AA862583 Mouse receptor tyr
28	133.5	6.4	863	19	AA862583 Mouse receptor tyr
29	133.5	6.4	863	19	AA862583 Mouse receptor tyr
30	133.5	6.4	863	19	AA862583 Mouse receptor tyr
31	132.5	6.3	869	18	AA862583 Mouse receptor tyr
32	132.5	6.3	869	18	AA862583 Mouse receptor tyr
33	131.5	6.3	537	22	AA877857 Human Dnk receptor
34	131.5	6.3	537	22	AA877857 Human Dnk receptor
35	131.5	6.3	537	22	AA877857 Human Dnk receptor
36	131.5	6.3	537	22	AA877857 Human Dnk receptor
37	130.5	6.2	1483	22	ABG16336 Human immunoglobul
38	128	6.1	592	23	AB815751 Human pancreas Gp3
39	128	6.1	592	23	AB815751 Human pancreas Gp3
40	128	6.1	708	23	AAU29315 Human PRO polypept
41	128	6.1	708	23	AAU29315 Human PRO polypept
42	126	6.0	1953	23	AA84351 Human gpi354 (putat
43	125.5	6.0	467	17	AA84094 Protein MYLX dilfe
44	123.5	5.9	340	22	AA80904 NSK2 receptor with
45	123.5	5.9	441	22	AAU0905 Human B lymphocyte

ALIGNMENTS

RESULT 1	AAU75541	standard; Protein; 386 AA.
ID	AAU75541	
AC	AAU75541	
DT	23-APR-2002	(first entry)
XX		
DE	Human B7-1-like protein, B7-1.h2.	
XX		
XX	Human: B7-1-like protein; B7-1; antiinfectility; gynaecological;	
KW	antitumour; cytostatic; immunosuppressive; antiarthritic; antineumatic;	
KW	antiinflammatory; dermatologic; antiproliferative; neuroprotective;	
KW	antidiabetic; haemostatic; antithyroid; antileuk; antiallergic;	
KW	antistimulant; nephrotoxic; antibacterial; virucide; tumour; cancer;	
KW	reproductive disorder; graft versus host disease; autoimmune disease;	
KW	toxic shock syndrome; allergy; nephropathy; skin disorder;	
KW	endocrinopathy; lymphoproliferative disorder.	
OS	Homo sapiens.	
XX		
PN	WO200200710-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	28-JUN-2001; 2001WO-US20719.	
XX		
PR	28-JUN-2000; 2000US-214512P.	
XX		
PA	28-NOV-2000; 2000US-0729264.	
XX		
PI	(AMGE-) AMGEN INC.	
	Welcher AA, Sarmiento UM, Schultz HJ, Chute HF;	

XX WPI: 2002-130881/17.
 DR N-PSDB; ABK13029.
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 XX
 PS Claim 13; Fig 2; 135pp; English.
 XX
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or all sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L-h2.
 XX
 S0 Sequence 386 AA;

Query Match 100.0%; Score 2088; DB 23; Length 386;
 Best Local Similarity 100.0%; Pred. No. 8.8e-165;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGAMENRPPSGSGNEVIEGPNARVYKSGQAFNCTVSGKLMMAISDMVLSV 60
 DB 1 MVAGAMENRPPSGSGNEVIEGPNARVYKSGQAFNCTVSGKLMMAISDMVLSV 60
 QY 61 RPEPIITNDRTSQRYYDGGNFTSEMIITHNPEPSDGINRCSLNSRLHSAVITLYVM 120
 DB 61 RPEPIITNDRTSQRYYDGGNFTSEMIITHNPEPSDGINRCSLNSRLHSAVITLYVM 120
 QY 121 GELFISVNLVVAENPCVYVLPSPHMTWLPDISWELDLVSHSSYFVPPSDLSQAVS 180
 DB 121 GELFISVNLVVAENPCVYVLPSPHMTWLPDISWELDLVSHSSYFVPPSDLSQAVS 180
 QY 181 ILALPQSGNLTCAVATKSLKARKSATVNLVIRCPDGTGGNINIGVSSLSLPSL 240
 DB 181 ILALPQSGNLTCAVATKSLKARKSATVNLVIRCPDGTGGNINIGVSSLSLPSL 240
 QY 241 PIMGVKVGGLAGTMTLTPCTLTTRCCCRRCGCCGCCPCRCRRKRGFRIOFGKS 300
 DB 241 PIMGVKVGGLAGTMTLTPCTLTTRCCCRRCGCCGCCPCRCRRKRGFRIOFGKS 300

QY 301 EKEKTKETETESGNSGNSDEKTDPTASLPKSCSSSPDEQRNCCPPHORADQR 360
 DB 301 EKEKTKETETESGNSGNSDEKTDPTASLPKSCSSSPDEQRNCCPPHORADQR 360
 QY 361 PPRPASHPOASFNLASPEKVSNTTVV 386
 DB 361 PPRPASHPOASFNLASPEKVSNTTVV 386
 RESULT 2
 ID AAU75540 standard; Protein; 382 AA.
 XX AAU75540;
 AC AAU75540;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Human B7-like protein, B7-L-h1.
 XX
 KW Human; B7-like protein; B7-L; antifertility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; antituber; antiallergic;
 KW antisthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PE 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX
 DR N-PSDB; ABK13028.
 DR
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 PS
 PS Claim 13; Fig 1; 135pp; English.
 XX
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as

CC Inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{h1}.

CC Sequence 382 AA;

Query Match 96.0%; Score 2004; DB 23; Length 382;
 Best Local Similarity 99.5%; Pred. No. 8e-158;
 Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GSGSGNEVIEGQPNARYLKGQARFNCVSGQWKLIMWALSDWVLSVRPMEPIITNDRF 72
 DB 9 GSGSGNEVIEGQPNARYLKGQARFNCVSGQWKLIMWALSDWVLSVRPMEPIITNDRF 68
 QY 73 TSQRVDGQGNFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTVQVWGLFIPSNLVY 132
 DB 69 TSQRVDGQGNFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTVQVWGLFIPSNLVY 128
 QY 133 AENEPCVTCPLPSHTWLPDISWELGLVSHSYFYFPEPSDLSAVSIILALTPQSNCTL 192
 DB 129 AENEPCVTCPLPSHTWLPDISWELGLVSHSYFYFPEPSDLSAVSIILALTPQSNCTL 188
 QY 193 TCVAATKSKARKSATVNLTVIRCPDPTGGGINTGVLSPSLGFSLPWKGVLGAG 252
 DB 189 TCVAATKSKARKSATVNLTVIRCPDPTGGGINTGVLSPSLGFSLPWKGVLGAG 248
 QY 253 TMLTPTCTLTIRCCCRRCGCCGCCGCCCRKRGRFROPKSKSEKTNKETETE 312
 DB 249 TMLTPTCTLTIRCCCRRCGCCGCCGCCCRKRGRFROPKSKSEKTNKETETE 308
 QY 313 SGNENSGYNSDEQKTDTSALPKSCSSDPEDQNSCCGPHQRADQRPAPSHPOASF 372
 DB 309 SGNENSGYNSDEQKTDTSALPKSCSSDPEDQNSCCGPHQRADQRPAPSHPOASF 368
 QY 373 NLASPEKVSNTTVY 386
 DB 369 NLASPEKVSNTTVY 382

RESULT 3
 AAU75542
 ID AAU75542 standard; Protein: 386 AA.

XX AAU75542;

XX 23-APR-2002 (first entry)

XX Human B7-1-like protein, B7-L_{h3}.

XX Human; B7-1-like protein; B7-L; anti-infectivity; gynaecological;
 XX antitumour; cytostatic; immunosuppressive; antirheumatic; antineoplastic;
 XX antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
 XX antidiabetic; nephrotropic; antibacterial; virucide; tumour; cancer;
 XX reproductive disorder; graft versus host disease; autoimmune disease;
 XX toxic shock syndrome; allergy; nephropathy; skin disorder;
 XX endocrinopathy; lymphoproliferative disorder.

XX Homo sapiens.
 OS
 XX WO200200710-A2.
 PN

XX 03-JAN-2002.
 PD 28-JUN-2001; 2001WO-US20719.
 XX 28-JUN-2001; 2000US-214512P.
 XX 28-JUN-2000; 2000US-214512P.
 PR 26-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 PA Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 PI WPI: 2002-130881/17.
 XX N-PSDB: ABK13030.
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis

Claim 13; Fig 3; 135pp; English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 The polypeptide, polynucleotide encoding it and antibody against (I) are
 useful for treating B7-like polypeptide-related disease, disorders or
 conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{h3}.

XX Sequence 386 AA;

Query Match 95.8%; Score 2000; DB 23; Length 386;
 Best Local Similarity 99.2%; Pred. No. 1.7e-157;
 Matches 371; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 GSGSGNEVIEGQPNARYLKGQARFNCVSGQWKLIMWALSDWVLSVRPMEPIITNDRF 72
 DB 13 GSGSGNEVIEGQPNARYLKGQARFNCVSGQWKLIMWALSDWVLSVRPMEPIITNDRF 72
 QY 73 TSQRVDGQGNFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTVQVWGLFIPSNLVY 132
 DB 73 TSQRVDGQGNFTSEMIHNVEPSDSGNIRCSLQNSRLHGSAYLTVQVWGLFIPSNLVY 132
 QY 133 AENEPCVTCPLPSHTWLPDISWELGLVSHSYFYFPEPSDLSAVSIILALTPQSNCTL 192

Db 133 AENPECEVTCLPSHWTRLPDISMELGLVSHSSYYFVPEPSDLSQASVSLATLPQSNGLT 192
 QY 193 TCVAATWKSARKSATYVNLVYICPPDPTGGGINIPGVLSLEPSLGFSTPTWKGVLGLAG 252
 Db 193 TCVAATWKSARKSATYVNLVYICPPDPTGGGINIPGVLSLEPSLGFSTPTWKGVLGLAG 252
 QY 253 TMLLPTCTLTIRCCGRRRCGCCGCCRRCKRREFRLOPKKSEKTKNETETE 312
 Db 253 TMLLPTCTLTIRCCGRRRCGCCGCCRRCKRREFRLOPKKSEKTKNETETE 312
 QY 313 SGNGNSGYNSDEKTTDTLSLPKSCSSDPQRRNSCGPPHQRADQPPRPAHQASF 372
 Db 313 SGNGNSGYNSDEKTTDTLSLPKSCSSDPQRRNSCGPPHQRADQPPRPAHQASF 372
 QY 373 NLASPEKVSNTYV 386
 Db 373 NLASPEKVSNTYV 386
 RESULT 4
 AAU75543
 ID AAU75543 standard; Protein: 377 AA.
 AC AAU75543;
 DT 23-APR-2002 (first entry)
 DE Human B7-1-like protein, B7-L_{H4}.
 KW Human: B7-1-like protein; B7-L; antiinfectivity; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antithyroid; antirheumatic;
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
 KW antidiabetic; nephrotropic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 OS Homo sapiens.
 PN WO200200710-A2.
 XX 03-JAN-2002.
 PD 28-JUN-2001; 2001MO-US20719.
 PE 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 PA Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
 PI WPI; 2002-130881/17.
 DR N-PSDB; ABK13031.
 XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 XX Claim 13; Fig 4; 135pp; English.
 PS The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L

CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{H4}.
 XX Sequence 377 AA;
 Query Match 75.5%; Score 1576; DB 23; Length 377;
 Best Local Similarity 88.6%; Pred. No. 2.4e-122;
 Matches 296; Conservative 12; Mismatches 20; Indels 6; Gaps 2;
 QY 1 MVAAMENRDPGGSGSNEVIEGPQNAARVKGSGQARFNCVSGWKLIMWALSDWVLSV 60
 Db 1 MVAAMENRDPGGSGSNEVIEGPQNAARVKGSGQARFNCVSGWKLIMWALSDWVLSV 60
 QY 61 RMEPIITNDRETSQRYDOGNGFTSEMIHNVEPSDGNIRSLQNSRLHGSAGVLYTVQV 120
 Db 61 RMEPIITNDRETSQRYDOGNGFTSEMIHNVEPSDGNIRSLQNSRLHGSAGVLYTVQV 120
 QY 121 GELFIPSVNLVAVNEPECEVTCLPSHWTRLPDISMELGLVSHSSYYFVPEPSDLSQASV 180
 Db 121 GELFIPSVNLVAVNEPECEVTCLPSHWTRLPDISMELGLVSHSSYYFVPEPSDLSQASV 180
 QY 181 ILALTPOSNGLTCVATWKSARKSATYVNLVYICPPDPTGGGINIPGVLSLEPSLGFSL 240
 Db 181 ILALTPOSNGLTCVATWKSARKSATYVNLVYICPPDPTGGGINIPGVLSLEPSLGFSL 240
 QY 241 PTWKGVLGLAGTMLLPTCTLTIRCCGRRRCGCCGCCRRCKRREFRLOPKKSEKTKNETETE 300
 Db 241 PTWKGVLGLAGTMLLPTCTLTIRCCGRRRCGCCGCCRRCKRREFRLOPKKSEKTKNETETE 300
 QY 301 EKEKTKETETESGNSGYNSDEKTTDTASLP 334
 Db 298 QTKKKQKVEKMT---PATIQNNKRRPQPLSLP 328
 RESULT 5
 ABG28169
 ID ABG28169 standard; Protein: 463 AA.
 AC ABG28169;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #28160.
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX

PN WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PR (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI, 2001-639362/73.
 DR N-PSDB; AAS92356.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT
 XX Claim 20; SEQ ID NO 58528; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 463 AA;

Query Match 63.1%; Score 1318; DB 22; Length 463;
 Best Local Similarity 99.6%; Pred. No. 7.6e-101;
 Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 49 MMALSDMVLVSRPMEPIITNDRTSORYDOGNFTSMITHNNEPSSGIRCSLNSR 108
 DB 1 MMALSDMVLVSRPMEPIITNDRTSORYDOGNFTSMITHNNEPSSGIRCSLNSR 60
 XX 109 LHGSATLTQVWGEFLIPSVNLVVAENPECEVTCPSHMTMLPDISELGLVSHSYFF 168
 DB 61 LHGSATLTQVWGEFLIPSVNLVVAENPECEVTCPSHMTMLPDISELGLVSHSYFF 120
 XX 169 VPEPSDQASVITALTPOSNGLTCAVATWKSLSKAKRSATVLTIRCPDPTGGGGINPG 228
 DB 121 VPEPSDQASVITALTPOSNGLTCAVATWKSLSKAKRSATVLTIRCPDPTGGGGINPG 180
 XX 229 VLSLPLGSLPTWKGKVGGLAGTMTLPTGCLTFRCCCRRCRCCGCCRCFCRR 288
 DB 181 VLSLPLGSLPTWKGKVGGLAGTMTLPTGCLTFRCCCRRCRCCGCCRCFCRR 240
 XX 289 KRG 291
 DB 241 KRG 243

RESULT 6

AAU75544
 ID AAU75544 standard; Protein; 370 AA.
 XX
 AC AAU75544;
 XX 23-APR-2002 (first entry)
 DE Mouse B7-1-like protein, B7-L_{ML}.
 XX
 XX Mouse; B7-1-like protein; B7-L; antiferility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antineumatic;
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; anticancer; antiallergic;
 KW antiaesthetic; nephrotoxic; antibacterial; virocidic; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 OS Mus musculus.
 PN WO200200710-A2.
 PD 03-JAN-2002.
 XX 28-JUN-2001; 2001WO-US20719.
 PF 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 PA Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
 PI WPI, 2002-130881/17.
 DR N-PSDB; ABR13032.
 XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 PT
 XX Claim 13; Fig 5; 135bp; English.
 XX The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC growth vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, inflammatory disease such as
 CC purpura and psoriasis, chronic inflammatory disease (Crohn's disease and ulcerative colitis),
 CC inflammatory bowel disease (Crohn's disease and diabetes mellitus). They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allensensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and

CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 XX sequence represents the amino acid sequence of mouse B7-1_{ML}.

SO Sequence 370 AA;

Query Match 43.8%; Score 913.5; DB 23; Length 370;
 Best Local Similarity 50.7%; Pred. No. 1.9e-67;
 Matches 192; Conservative 50; Mismatches 102; Indels 35; Gaps 5;

14 SSGSGNVEIEGPONARVYKSGQARENCTVSQGKLMALSDMVYLSVPMPTITNDRE 73
 DB 21 SSSSIIIEGPONVYVLDKSEAHFNCTVHGKMLMTLNQMVYLSLTQGPITITNRR 80
 QY 74 SGRYDGGNFTSEMTIHNVPEPSDGNICSLQNSRLHSAVLYQVWGLFTPSVNLVA 133
 DB 81 YASYNSTDSFSELIHNVQPSDGSVQCSLQNSHGFSAFLSVQVWGLTINPSNNLI 140
 QY 134 EMEPEVTCPSHWYLPDISEWELGLVSHSSYFVPEPSDLSAVSIIALTFQSGNTIT 193
 DB 141 EEPENVTICVAVGWTSLPDISMELEVPVSHSSYNSFLEPNFMYVLSVLDITLGNTIT 200
 QY 194 CVATWKSLEKRSATVNLVTRCPDGGGINIPGVLSLPSLGFSLPTWKGKGLIAGT 253
 DB 201 CVAEKLDLQASKSLVNLVTVQPPD-----STGEGPALPTWAILLAVAFS 248
 QY 254 MLPLPTGTRCCGCCRCGCCGCCRCGCRKRRGRFIDPQ--KKESEKTKR--- 307
 DB 249 LLLIILVLIITFCG-----CASRREKESFYQNIIRKSAMRNKADP 293
 QY 308 ETESENGENSGYNSDOKTTPDASLPKSCSESDPEQNSGCPRHQADRPAPASH 367
 DB 294 ETKRSKEKENVYSSDEAKAQAQASLPKSAEVSLEPKRSSSL--PYQELNKHQGPAT 351
 QY 368 PQASFNLASEKYSNTTV 386
 DB 352 PRVSFLASPKVKNVTLV 370

RESULT 7
 AAU75547
 ID AAU75547 standard; Protein; 631 AA.
 AC AAU75547;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Rat B7-1-like protein, B7-1.

XX Rat; B7-1-like protein; B7-1; antiinfectivity; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KW antidiabetic; dermatological; antipsoriatic; neuroprotective;
 KW antitubercular; haemostatic; antithyroid; anticancer; antiallergic;
 KW antiasmatic; nephrotropic; antibacterial; virocid; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 XX endocrinopathy; lymphoproliferative disorder.

OS Ratus rattus.

XX MO200200710-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001MO-US20719.

XX 28-JUN-2000; 2000US-214512P.

XX 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.

XX

PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT.
 XX WPI; 2002-130881/17.

PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 XX proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Disclosure: Fig 8; 135pp; English.

CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of rat B7-1.

XX
 SO Sequence 631 AA;

Query Match 27.4%; Score 572.5; DB 23; Length 631;
 Best Local Similarity 27.8%; Pred. No. 7.1e-39;
 Matches 163; Conservative 44; Mismatches 113; Indels 267; Gaps 9;

QY 14 SSGSGNVEIEGPONARVYKSGQARENCTVSQGKLMALSDMVYLSVPMPTITNDRE 72
 DB 69 SGIYGVYKSKRVNTVYVLDKSEAHFNCTVHGKMLMTLNQMVYLSLTQGPITITNRR 128
 QY 73 -----TSQRDQGNFTS 85
 DB 129 KALSCDYKFCSEOSIHTIYQKDKMVLVLSGVPEVWPKYKKNRTYASTNSDPSIS 188
 QY 86 EMITHNVPEPSDGNICSLQNSRLHSAVL----- 115
 DB 189 ELIITHVQPSDGSVQCSLQNSHGFSAFLSVQVYDIANNYSFSLGLILSDRGTYTCV 248
 QY 116 -----TVQWGELEFIPSVNLVYVWENPCVYTCPSHWYLPDISWELGLV 161
 DB 249 QREYGGSYVYKHLTFYVWGLTINPSNNLIYEGPCNVTCYAVGWTSLPDISWELV 308
 QY 162 SHS-----SYFVPEPS 173
 DB 309 SHSLSVRADPPTNITEYGNPSADIKRITCFASGGFPKPRLSWLENGRELNSYNSFLEPG 368

QY 174 DLOSASLALTPQSNGLTFCVATWKSILKARKSATVNL----- 211
 Db 369 NEMRLVSLDLTPANGNLTFCVAKELDLQASKSLTNVNLGINTTISDPESELYTSSOLD 428
 QY 212 -----TVTRCPDGTGGGINIPGVLSLPSLGFSLPTWKG 247
 Db 429 FNATYDFHIDCFIEYGAHVSQNFVVOPPD-----SIGEGPALPTWAIL 476
 QY 248 LGIAGTML----- 256
 Db 477 LAVAFSLILLIIVLIIFTWKPPEDPDEKQTVFPAWAGPDVAKIIFIAITVAV 536
 QY 257 TPTCTLTTRCCCGRRCCGCCGCCCKRRKGFRIQFO--KSEKTKTK---ETETE 312
 Db 537 IAAIILIIFFCCCAAR-----REKESTYQNEIRKSANNRINKADPETK 582
 QY 313 SGENSGNSDQKTTDASLPPKSCSSDDEQRNSCGPPHORADQ 359
 Db 583 SEKENYGSSEDCITVKFRRCRRRNRNSRETNKMLYIGPVEAAEQ 629

RESULT 8

AAU75545
 ID AAU75545 standard: Protein: 270 AA.

AAU75545:

23-APR-2002 (first entry)

Mouse B7-1like protein, B7-L_{m2}.

Mouse; B7-1like protein; B7-L; antifertility; gynaecological;
 antitumor; cytostatic; immunosuppressive; antirheumatic;
 antiinflammatory; dermatological; antiproliferative; neuroprotective;
 antidiabetic; haemostatic; antihypertensive; antiallergic;
 antiproliferative; nephrotoxic; antibacterial; virucide; cancer;
 reproductive disorder; graft versus host disease; autoimmune disease;
 toxic shock syndrome; allergy; nephropathy; skin disorder;
 endocrinopathy; lymphoproliferative disorder.

Mus musculus.

MO200200710-A2.

03-JAN-2002.

28-JUN-2001; 2001WO-US20719.

28-JUN-2000; 2000US-214512P.

28-NOV-2000; 2000US-0729264.

(AMGE-) AMGEN INC.

Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;

WPI; 2002-130881/17.

N-PSDB; ABK13033.

New B7-like polypeptides, polynucleotides and their modulators, useful
 for diagnosing, preventing and treating reproductive, immune and
 proliferative disorders, e.g. cancer and arteriosclerosis

Claim 13; Fig 6; 135pp; English.

The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 The polypeptide, polynucleotide encoding it and antibody against (I) are
 useful for treating B7-like polypeptide-related disease, disorders or
 conditions including reproductive disorders (e.g. infertility
 miscarriage, preterm labour and delivery and endometriosis) and
 proliferative disorders. Antibodies, soluble proteins comprising
 extracellular domains and other regulators of B7-L polypeptides are
 useful for enhancing the immune response to tumours. (I) plays a role in
 growth and maintenance of cancer cells based on the observation of

seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 polypeptide. Hence modulators of (I) are useful for the treatment of
 cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 in allograft transplantation, graft versus host disease, T-cell
 dependent B-cell mediated diseases and autoimmune diseases. B7-L
 molecules are useful for alleviating the symptoms associated with
 diseases involving chronic immune cell dysfunction or to treat
 autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 purpura and psoriasis, chronic inflammatory disease such as
 inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 are also useful as immunosuppressive agents for bone marrow and organ
 transplantation or to prolong graft survival. B7-L molecules are also
 useful for diagnosis and treatment of diseases involving abnormal cell
 proliferation, including arteriosclerosis and vascular restenosis.
 Antagonists of B7-L polypeptides are useful for alleviation of toxic
 shock syndrome or allodensitisation due to blood transfusions, and for
 treatment of allergy, asthma and hypersensitivity reactions,
 nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 lymphoproliferative disorders such as multiple myeloma. The present
 sequence represents the amino acid sequence of mouse B7-L_{m2}.

Sequence 270 AA:

Query Match 27.1%; Score 566.5; DB 23; Length 270;
 Best local similarity 34.6%; Pred. No. 7.7e-39;
 Matches 131; Conservative 42; Mismatches 71; Indels 135; Gaps 5;

QY 14 SSGNEVEIEGPNRNVKLGSGARFCTVSQGLIMNALSDMVVLSVRPMEPIITNDRET 73
 Db 21 SSSYQIIEGQNTVYVKDSEAHNCTYTHGKMLMTNLQMVVLSFTQGPITNNRFT 80
 QY 74 SQRDQGNFTSEMTIHNVPSDGNTRCSLQNSRLGASLTLYQWGELEFPVNLVVA 133
 Db 81 YASNSTDSFISLIIHDVQSDSGVQSLQNSHGFSAFLVQV----- 125
 QY 134 ENRPEVTCPSHMTWLDISWELGLVSHSSYFVPEPDSLOSASLIALTPQSNGLT 193
 Db 126 ----- 125
 QY 194 CVATWKSILKARKSATVNLTVTRCPDGTGGGINIPGVLSLPSLGFSLPTWKGVLGAGT 253
 Db 126 -----DSIGEGPALPTWAILLAVAFS 148
 QY 254 MLTPPTCTLTCCCGRRCCGCCGCCCKRRKGFRIQFO---KSEKTKTK--- 307
 Db 149 LLLIILVLIIFCC-----CASRKEESTYQNEIRKSANNRINKADP 193
 QY 308 ETEESGNSGNSDQKTTDASLPPKSCSSDDEQRNSCGPPHORADOPRPAASH 367
 Db 194 ETKLKGKRNKYGVSDPAKAQAOTASLPPKSAVSLPEKSSSL--PYDELNKHQGPAT 251
 QY 368 PQASENLASPEKVSNTTV 386
 Db 252 PRVSFDIASPQKRVNNTLV 270

RESULT 9

AAU75546
 ID AAU75546 standard: Protein: 223 AA.

AAU75546:

23-APR-2002 (first entry)

Mouse B7-1like protein, B7-L_{m3}.

KM Mouse; B7-1-like protein; B7-L; antiinfectivity; gynaecological;
 KM antitumor; cycostatic; immunosuppressive; antiarthritic; antineumatic;
 KM antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antileuc; antiallergic;
 KM antistatic; nephrotoxic; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; nephropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder.
 OS Mus musculus.
 XX WO200200710-A2.
 XX 03-JAN-2002.
 XX 28-JUN-2001; 2001WO-US20719.
 XX 28-JUN-2000; 2000US-214512P.
 XX 28-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 PA Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
 PI MPI; 2002-130881/17.
 DR N-PSDB; ABR13034.
 XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 PS Claim 13; Fig 7; 135pp; English.
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allosteric sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Graves' disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m3}.
 XX
 SX Sequence 223 AA;
 Query Match 23.2%; Score 485; DB 23; Length 223;

Best Local Similarity 31.4%; Pred. No. 3.4e-32;
 Matches 118; Conservative 37; Mismatches 45; Indels 176; Gaps 6;
 QY 14 SSGNEVEIEGPQNAVLEKGSQAREFCVSGMKLIMNALSQVLYSRPHEPITITNREF 73
 DB 21 SSSVQIIEGQNAVYVLEKDSAHFNCTVTHGKMLIMNALSQVLYSRPHEPITITNREF 80
 QY 74 SORDGQNEFTSEMTHNVEPSDGNIRCSLQNSRLHGSAYLVQVWGELEFIPSNLVVA 133
 DB 81 YASYNSTDSFISELIIHDVOPSDSGSVQCSLQNSHGRGSAFLSVQ----- 125
 QY 134 ENPECEVTCPLPSHMTWLPDISMELGLVSHSSYFVEPEPSDGSAYITALTPQSNCTLT 193
 DB 126 -----ESTY-----ONEI----- 133
 QY 194 CVATWKSILKARKSATVNLTVIRCPDGTGGGINIPGVLSLPSLGFSLPTWKGVLGAGT 253
 DB 134 -----RKSANM----- 139
 QY 254 MLNPTCTLTIRCCGCCRCGCCGCCRCPCCKRKRGRFIOFKSEKERTNK---ETE 310
 DB 140 -----RTMKADPETK 149
 QY 311 TSGNENSGYNSDEOKTTDTASLPKSCSESSDPEORNSCGPPHORADOPPPAPASHQA 370
 DB 150 LKSGKENYGVSSDEAKAQTASLPKSAEVSLEPKRSSL--PYQELNKHQEPATHPRV 207
 QY 371 SFNIASEKXSNTTVV 386
 DB 208 SFDIASPKVKNVTLV 223
 RESULT 10
 AAB81925
 ID AAB81925 standard; protein; 404 AA.
 AC AAB81925;
 XX 15-JUN-2001 (first entry)
 DT Extracorporeal circulation material receptor protein.
 DE Extracorporeal circulation material receptor protein.
 KW Extracorporeal circulation; carbonyl stress product; receptor;
 KM diabetes; vascular lesion; excretory dysfunction.
 OS Unidentified.
 XX WO200118060-A1.
 PN 15-MAR-2001.
 PD 08-SEP-2000; 2000WO-JP06172.
 PF 08-SEP-1999; 99JP-0254463.
 PR (TORA) TORAY IND INC.
 PA Shimizu S, Kubota M, Akiyama H, Usui M;
 PI MPI; 2001-290314/30.
 DR Material for extracorporeal circulation, applicable in selective
 XX elimination of diabetic complication factors such as carbonyl stress
 XX products caused by abnormally promoted carbonyl stress from excretory
 XX dysfunction in vascular lesions -
 PS Claim 1; Page 31-32; 36pp; Japanese.
 XX The present invention describes a material for extracorporeal circulation
 CC which is made from a water-insoluble carrier immobilized with a protein
 CC having the sequence shown here. The materials of the invention, including
 CC adsorbents, are for extracorporeal circulation, which are applicable in
 CC the selective elimination of diabetic complication factors from a body

fluid, and are therefore useful in treating vascular lesions like
arteriosclerosis due to carbonyl stress products caused by abnormally
promoted carbonyl stress from excretory dysfunction.

Sequence	404 AA
SQ	

Query Match	7.0%;	Score 145.5;	DB 22;	Length 404;
Best Local Similarity	23.5%;	Pred. No. 0.001;		
Matches 77;	Conservative 34;	Mismatches 94;	Indels 123;	Gaps 15

```

OY 64 BEITNDNFTS-----ORIOGGNET--SMITHNEPDSGNIR---CSLONSMLHS 11
Db 162 KLVNNEGVSVKREOTRHHPTGTLTQSELM---VTPANGDPRPFSCSFPGLPRRR 218
OY 113 AYLTVOVNGELEIP---SYNLVVAENP-----CEYTCJLPSHWTWLPDJS 154
Db 219 ALRTAPIQRPVPEPBLEEVLV---EPREGAVAPGVTVLTCEVPAQS-----FOH 270
OY 155 WELGILVSHSSYFYFPEPSDQSAVSLALTPQSNGLTLCVATWKSIAKNSATVNLTY 214
Db 271 WKMD-----GVPLPLPSPVLTILPEIPPOQGYSCVATHSHSGPESRAVSIST 321
OY 215 RCPQDTGGGINIPVULSLPLGFLPLPWGVLGAGTMTLT-----PTCTLTIRC 267
Db 322 E-PGEEG-----PLAGSVGGSGGLTALALGILGGTALLGVI 361
OY 268 CRRRCGCCGCCGCCFCFCRRRRGRFIOFOKKSEKKT--KKEFTESGNMNSYNSDQ 325
Db 362 LMQRR-----QRGGERKAPENOEDEBERALELN-----389
OY 326 KTTDTLASLPKSCESDPPORNSCGPP 353
Db 390 -----QSEPEPEGESSTGPF 404

```

RESULT 11
AAE23219
ID AAE23219 standard; Protein; 404 AA.

AAE23219

DT 27-AUG-2002 (first entry)

XX
ne Human receptor for advanced glycosylation end product (RAGE) protein.

Human; Receptor for advanced glycosylation end product; RAGE; cardiomyocyte; tissue growth; neointimal formation; blood vessel; restenosis; diabetes; myocardial infarction; angioplasty; peripheral vascular surgery; angina; transcatheter aortic; acute thrombotic stroke; venous thrombosis.

OS Homo sapiens.

PN WO200230889-A2

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US32036

PR 13-OCT-2000; 2000US-0687528

PA (UYCO) UNIV COLUMBIA NEW YORK

PI Stern DM, Schmidt A, Maiso S, Lopez D, Sanchez M

DR WPI; 2002-426260/43
N-PSDB: AAD36952

XX
XX
BT

XX Inhibiting new tissue growth or neointimal formation in blood vessels
PT of subject suffering from diabetes, stroke and preventing restenosis,
PT comprises administering inhibitor of receptor for advance glycation end
PT product -

PS Disclosure; Page 16; 43pp; English

XX The invention relates to a method for inhibiting new tissue growth or
CC neointimal formation in blood vessels in a subject that has experienced
CC blood vessel injury and preventing exaggerated restenosis in a diabetic
CC subject. The method comprises administering an inhibitor of receptor for
CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new
CC tissue growth or neointimal formation in subject's blood vessels and
CC preventing restenosis in the subject. The method is useful for inhibiting
CC new tissue growth or neointimal formation in blood vessels in a subject
CC like non-human animal, a transgenic non-human animal or a human suffering
CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable
CC angina, myocardial infarction, abrupt closure following angioplasty or
CC stent placement, or thrombosis as a result of peripheral vascular surgery
CC The method is also useful for preventing restenosis and for determining
CC whether a compound inhibits new tissue growth in a blood vessel in a
CC subject. The present sequence is human receptor for advanced
CC glycosylation end product (RAGE) protein.

Sequence	404 AA
SQ	

Query Match	7.0%;	Score 145.5;	DB 23;	Length 404;
Best Local Similarity	23.5%;	Pred. No. 0.001;		
Matches	77;	Conservative 34;	Mismatches 94;	Indels 123;
				Gaps 15;

```

0Y      64  EPIITNDRETS-----QRXDOGANTF--SEMI IHVSEDSGNTF--OSLOMSKTHS 112
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      162 KPLPNEKGVSAKEOTRRPPEGLFTLOSELM--VTPARGDPRPFSCFSFSGLEPRHR 218
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y      113 AYLTVOMGELFIP--SVNLVAENEP-----CEVYCLPHTMTNLPDIS 154
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      219 ALRNPAPLOPMEVPEVPLEEVOLV--EPEGGAVALPGGTVITLCVPAPPS-----PQIH 270
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y      155 WELGLVSHSHSYFEVPEPDSLOASVILALFQSGNTLCTVATWMSLKARKASVNLTVI 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      271 WMKD-----GVPLPLPSPVULLPELIGPODQCTYSCAVTHSHGPOESRAVSII 322
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y      215 RCPDITGGGINIPGVLSLPSLCSLPTWNGKVLGAGTMMLT-----PCTVITIRC 267
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      322 E-PEEG-----PTASVSGSLGTLALALGLSGTALALIGVI 361
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y      268 CCRRCGCCNCCRCFCRCRRKRGFRIOFOKSEKERT--NKETETESGENSGVNSDQ 322
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      362 LMQR-----ORREKRAPENOEEERAEALN----- 38
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y      326 KTTDTASLPPKSCBSDDPEOBNSSGPP 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      390 -----OSEPEAGESSTGGP 404
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12
AAU77543
ID AAU77543 standard; Protein; 404 AA.

AAU77543

05-JUN-2002 (first entry)

Human receptor for advanced glycosylation end product (HAGE)

XX Receptor for advanced glycation end product; RAGE; receptor;
KW amyloid beta peptide; blood-brain barrier; neurovascular stress;
KW amyloid beta peptide; blood-brain barrier; neurovascular stress;
KW cerebral vasoconstriction suppressor; cerebral blood flow enhancer;
KW cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor
KW cerebral amyloid angiopathy; head trauma; stroke; human
KW disease; Down's syndrome

OS Homo sapiens.

PN WO200214519-A1

PD 21-FEB-2002

PF 14-AUG-2001; 2001WO-0523410

PR 14-AUG-2000; 2000US-0638648.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 XX Stern DM, Schmidt AM, Yan SD, Zlokovic B;
 PI
 DR WPI; 2002-257610/30.
 DR N-PSDB; ABK10856.

PT Ameliorating neurovascular stress and decreasing cerebral
 PT vasoconstriction in subject suffering from chronic/acute cerebral
 PT amyloid angiopathy, by administering inhibitor of receptor for advanced
 PT glycation endproduct
 XX
 PS Disclosure; Page 16; 68pp; English.

CC The invention describes a method of ameliorating neurovascular stress,
 CC and decreasing cerebral vasoconstriction in subject suffering from
 CC chronic or acute cerebral amyloid angiopathy, comprising administering
 CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).
 CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain
 CC barrier, thus decreasing cerebral vasoconstriction and increasing
 CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
 CC subject, decreasing cerebral vasoconstriction in a transgenic non-human
 CC animal (preferably, transgenic mouse overexpressing mutant human amyloid
 CC beta precursor protein) or a human, suffering from chronic or acute
 CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
 CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy
 CC in a subject, where the neurovascular stress is caused by Alzheimer's
 CC disease, aging, Down's syndrome, head trauma or stroke. This is the
 CC amino acid sequence of human receptor for advanced glycation end
 CC product (RAGE) described in the invention.

XX Sequence 404 AA;

Query Match 7.0%; Score 145.5; DB 23; Length 404;
 Best Local Similarity 23.5%; Pred. No. 0.001;
 Matches 77; Conservative 34; Mismatches 94; Indels 123; Gaps 15;

QY 64 EPIITNDRTS-----QRYDGGNFT--SEMIHNVEPSDSGNIR-----CSLQNSRLHGS 112
 DB 162 KPLVFNKGVSVKQETRRHHPETGLTQSELM---VTPARGDPRPTFSCSFSPGLPRHR 218
 QY 113 AYLTVQVMGELFIP-----SVNLVVAENP-----CEVTCLPSHMTWLPDIS 154
 DB 219 ALRTAPIQPRVWEPEVLEEVQVYV---EPEGGAVAPGTVTLTCEVPAQPS-----PQIH 270
 QY 155 WELGLLVSHSYTFPEPSDLSAVSIALTPQSGTLTCVATWMSLKARSAVNLTVI 214
 DB 271 WMKD-----GVPLPLPSPVILPEIGPDQGTSCVATHSSHGQESRAVSISII 321
 QY 215 RCPDITGGGINIPGVLSLPSLGFSLPTWKGVLGLAGTMLT-----PTCTLTTRCC 267
 DB 322 E-PGEGG-----PTAGSVGSGGLTALALGILGIGTALLIGVI 361
 QY 268 CRRRCGCCGCCRCRCCRRKRGFRIOFQKSEKERT--NKETETESGNGNSGYNSDEQ 325
 DB 362 LMQR-----GVPLPLPSPVILPEIGPDQGTSCVATHSSHGQESRAVSISII 389
 QY 326 KTTDTASLPKSCSSDPEQNRSSCGPP 353
 DB 390 -----QSEEPAGESSTGPP 404

RESULT 13
 AAM48745
 ID AAM48745 standard; protein; 404 AA.
 XX
 AC AAM48745;
 XX
 DT 02-APR-2002 (first entry)
 XX
 DE Human RAGE protein SEQ ID NO 1.

XX Human: RAGE; receptor for advanced glyated endproduct; receptor;
 KW antidiabetic; neuroprotective; cytostatic; antiinflammatory; vasotropic;
 KW nephrotropic; dermatological; antiarteriosclerotic; nootropic; diabetes;
 KW Alzheimer's disease; cancer; inflammation; kidney failure;
 KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200192892-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US17447.
 XX
 PR 30-MAY-2000; 2000US-2073437.
 PR 05-MAR-2001; 2001US-0799152.
 XX
 PA (TRAN-) TRANS TECH PHARMA.
 XX
 PI Shabbaz M;
 DR WPI; 2002-114372/15.

PT Detecting a receptor for advanced glyated endproducts (RAGE)
 PT modulators, for treating e.g., cancer, diabetes or inflammation,
 PT comprises measuring the amount of bound anti-RAGE antibody
 XX
 PS Claim 1; Fig 2; 49pp; English.

CC The invention relates to detecting receptor for advanced glyated
 CC endproducts (RAGE) modulators comprises determining the amount of RAGE
 CC protein or its fragment bound to the pre-adsorbed ligand by measuring the
 CC amount of anti-RAGE antibody bound to the solid surface. The method is
 CC useful for rapid, high-throughput identification of compounds that
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis
 CC or inflammatory lupus nephritis, erectile dysfunction and
 CC atherosclerosis.

XX Sequence 404 AA;

Query Match 7.0%; Score 145.5; DB 23; Length 404;
 Best Local Similarity 23.5%; Pred. No. 0.001;
 Matches 77; Conservative 34; Mismatches 94; Indels 123; Gaps 15;

QY 64 EPIITNDRTS-----QRYDGGNFT--SEMIHNVEPSDSGNIR-----CSLQNSRLHGS 112
 DB 162 KPLVFNKGVSVKQETRRHHPETGLTQSELM---VTPARGDPRPTFSCSFSPGLPRHR 218
 QY 113 AYLTVQVMGELFIP-----SVNLVVAENP-----CEVTCLPSHMTWLPDIS 154
 DB 219 ALRTAPIQPRVWEPEVLEEVQVYV---EPEGGAVAPGTVTLTCEVPAQPS-----PQIH 270
 QY 155 WELGLLVSHSYTFPEPSDLSAVSIALTPQSGTLTCVATWMSLKARSAVNLTVI 214
 DB 271 WMKD-----GVPLPLPSPVILPEIGPDQGTSCVATHSSHGQESRAVSISII 321
 QY 215 RCPDITGGGINIPGVLSLPSLGFSLPTWKGVLGLAGTMLT-----PTCTLTTRCC 267
 DB 322 E-PGEGG-----PTAGSVGSGGLTALALGILGIGTALLIGVI 361
 QY 268 CRRRCGCCGCCRCRCCRRKRGFRIOFQKSEKERT--NKETETESGNGNSGYNSDEQ 325
 DB 362 LMQR-----GVPLPLPSPVILPEIGPDQGTSCVATHSSHGQESRAVSISII 389
 QY 326 KTTDTASLPKSCSSDPEQNRSSCGPP 353
 DB 390 -----QSEEPAGESSTGPP 404

RESULT 14

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:24:32 ; Search time 37.1283 Seconds
(without alignments)
2142.147 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077
Sequence: 1 MERHLLTPEAVGSGSGNFV.....HPQASFNLASPEKVSNTTV 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_dactilep:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1694	81.6	315	4 Q9NS15	Q9NS15 homo sapien
2	579.5	27.9	270	11 Q9D8G2	Q9D8G2 mus musculu
3	142.5	6.9	344	13 Q9DF61	Q9DF61 gallus gall
4	141	6.8	168	11 Q9D912	Q9D912 mus musculu
5	137.5	6.6	315	13 Q9DG15	Q9DG15 gallus gall
6	137.5	6.6	344	11 Q9DPJ0	Q9DPJ0 mus musculu
7	137.5	6.6	344	13 Q93242	Q93242 gallus gall
8	136.5	6.6	173	11 Q9D4K2	Q9D4K2 mus musculu
9	135	6.5	1496	4 Q92626	Q92626 homo sapien
10	133.5	6.4	344	4 Q9P121	Q9P121 homo sapien
11	133	6.4	1427	13 Q91562	Q91562 xenopus lae
12	131	6.3	697	11 Q9DBP0	Q9DBP0 mus musculu
13	131	6.3	697	11 Q92290	Q92290 mus musculu
14	129.5	6.2	164	5 Q22048	Q22048 caenorhabdi
15	129	6.2	6632	5 Q01761	Q01761 caenorhabdi
16	128	6.2	1445	11 Q63155	Q63155 rattus norv

17	127.5	6.1	197	5 Q17641	Q17641 caenorhabdi
18	127.5	6.1	869	4 Q15146	Q15146 homo sapien
19	127.5	6.1	871	11 Q61987	Q61987 mus musculu
20	127.5	6.1	881	11 Q61988	Q61988 mus musculu
21	127	6.1	313	13 Q57596	Q57596 gallus gall
22	126	6.1	6620	4 Q96AA2	Q96AA2 homo sapien
23	125.5	6.0	166	5 Q950Y1	Q950Y1 caenorhabdi
24	125.5	5.9	188	5 Q18238	Q18238 caenorhabdi
25	122.5	5.9	455	4 Q9U1R0	Q9U1R0 homo sapien
26	122.5	5.9	4370	4 Q9H3V5	Q9H3V5 homo sapien
27	120	5.8	695	11 Q9J109	Q9J109 rattus norv
28	120	5.8	959	5 Q9N9Y9	Q9N9Y9 drosophila
29	120	5.8	968	5 Q9W4T9	Q9W4T9 drosophila
30	120	5.8	975	5 Q97174	Q97174 drosophila
31	119.5	5.8	6632	5 Q17362	Q17362 caenorhabdi
32	117.5	5.7	390	4 Q9H1X9	Q9H1X9 homo sapien
33	117.5	5.7	5636	4 Q96RW7	Q96RW7 homo sapien
34	116.5	5.6	345	5 Q9W4U1	Q9W4U1 drosophila
35	116	5.6	225	11 Q9UK39	Q9UK39 mus musculu
36	116	5.6	2221	5 Q9U1M1	Q9U1M1 drosophila
37	115.5	5.6	1596	4 Q9HCL6	Q9HCL6 homo sapien
38	114.5	5.5	46	4 Q9XVX3	Q9XVX3 caenorhabdi
39	114	5.5	569	11 Q920G3	Q920G3 mus musculu
40	114	5.5	868	11 Q62838	Q62838 rattus norv
41	114	5.5	1395	5 Q44924	Q44924 drosophila
42	114	5.5	1395	5 Q9W213	Q9W213 drosophila
43	114	5.5	2673	4 Q96SC3	Q96SC3 homo sapien
44	114	5.5	196	4 Q9H0T1	Q9H0T1 homo sapien
45	113.5	5.5			

ALIGNMENTS

RESULT 1

ID	Q9NS15	PRELIMINARY:	PRT:	315 AA.
AC	Q9NS15			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	IGSF5 protein (Fragment).			
GN	IGSF5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,			
RA	Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,			
RA	Takagi T., Sakaki Y., Taudien S., Bleeschmidt K., Pollay D.,			
RA	Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,			
RA	Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,			
RA	Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,			
RA	Shinomiya S., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,			
RA	Mitsushima S., Shimizu N., Nordick G., Hornischer K., Brandt P.,			
RA	Ramsfer J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,			
RA	Wehrmeyer S., Borzym K., Gardner K., Nizetic D., Francis F.,			
RU	Submitted (May-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AL163280; CAB90447.1; -			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR003599; Ig.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	Pfam: PF00047; Ig_2.			
DR	SMART: SM00409; Ig_2.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE			

[illegible]

Query Match 6.8%; Score 14; DB 11; Length 168;
 Best local Similarity 32.0%; Pred. No. 7/e-05;
 Matches 39; Conservative 20; Mismatches 49; Indels 14; Gaps 6;

260 CTTTCCCGCCGCGC-CGCGCCGRRRGPR---IOFGSKSEKTKNETETESG 314
 55 CTSLSRCCGCGCCGKC-CHGCGCGC--CCSRKSRFRSKTTLKFPQITEKGEGSLQRRTRRG 111

```

Query Match      6.6%; Score 137.5; DB 13; Length 315;
Best Local Similarity 26.0%; Pred. No. 0.00036;
Matches          70; Conservative 38; Mismatches 104; Indels 57; Gaps 14.

QY      3 RHLLTVEAVGSGSGNEVI-EGPONATVYLKGSQARENCITYSQ-----GW---KLIMAPLS 53
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB     20 RLFLTVFAGVYFRFGDATTFRKAMDNVTYVRQGESATLRLCSVDNRTRYAAMLNRSSILTAGN 79
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY      54 DMVLVSRPMPKITTNDRFTSORRDQGNFTSEMIHNVPEPSDSGINRSLQ-----NS 107
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB     80 DKWCLDPFPVLLANTKTQYSIQ-----INDVYDEGGPYTCGVQTDNHPKTS 126
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY     108 RLHGSAVLTVOVMGEPLTFPSYNLVVAENEPCEVTLCPSHMTRLPD--ISWELGLVSHSS 165
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB    127 RVH----LIYQVSKRIETISSDISINEGANVSLCIA--TGREDPIITWR--HISPKA 176
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY     166 YEVPEPSDLQASVIALTPQSNGTLTCVATWMSLKARSAFYNLIV-----IRCPQDT 220
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB    177 VGFISEDEYLE---ITGITREQSGEVECHS--NDVAAPVVQRVKVINYNPPTISDAKST 231
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY     221 GGGINIPGVL---SLSPILGFSLEPTWKC 245
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB    232 GPVPQGKGILMCEASAVPSADFQ---WK 257
           |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 6
Q99PJ0 PRELIMINARY: PRT: 344 AA.
ID Q99PJ0
AC Q99PJ0;
DF 01-JUN-2004 (TREMBLrel. 17, Created)
```

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Neurotrophin.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=BRAIN;
 RA Kim T.H., Choi S.C., Kim J., Jeon J.W., Kim K.D., Lee S.H.;
 RT "Cloning and expression of mouse neurotrophin gene in the developing
 RT nervous system".
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF282980; AAK0276.1;
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003600; Ig_1ike.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00409; IgC2; 3.
 DR SMART: SM00410; Ig_1ike; 2.
 KW Immunoglobulin domain.
 SQ SEQUENCE 344 AA; 37924 MW; 3ECC6D5EE6C5C17D CRC64;
 Query Match 6.6%; Score 137.5; DB 11; Length 344;
 Best Local Similarity 26.0%; Pred. No. 0.0004;
 Matches 68; Conservative 36; Mismatches 115; Indels 43; Gaps 13;
 QY 3 RHLLTPEAVGSSGSGNEVI-EGPQNAVTLKSGQARFNCYSGQKLMALSLMVLVSR 61
 DB 20 RLFLVTPGVAVRSGDTPFKAMDNTVROGESATLRCSDVNRVAVLNRSLTYGN 74
 QY 62 PMEPIITNDRF-TSQRYDQGNFTSE--MIHNVESDSGNGINCSLO-----NSRLHGS 112
 DB 75 ---PYAGNDKWCIDPRVLLIGNTQTOYSIEIQNVVDYDESPYCSQVQIDNHPTSVH-- 129
 QY 113 AYLTVQMGELFIPSVNLVVAENPECEVTCLPSSHTRLPD--ISWELGLVSHS 172
 DB 130 --LIVQSPKITEISSISINEGNNISLFCIAIGRPE-PLTVWR--HISPAVGFVSD 183
 QY 173 SDIQSAVSLATLPQSNGLTLCVATWKSILKARKSATVNLTVICP-----ODTGGGINP 227
 DB 184 EYQD-----IOGTIREQSGEYECAS-NDVAAPVQVRKVTAVNPPIFEAKGTGVPVGOK 238
 QY 228 GVL-----SLPSLGFSLPTWCK 245
 DB 239 GTLQCEASAVPSAEQO---WFK 257
 RESULT 7
 ID 093242 PRELIMINARY; PRT; 344 AA.
 AC 093242:
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CEP-1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_Taxid=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimura Y., Shirabe K., Fukushima M., Takeshita M., Tanaka H.;
 RT "CEP-1: an immunoglobulin superfamily molecule, Has Cell Adhesion
 RT Activity and Shows Dynamic Expression Patterns in Chick Embryonic
 RT Spinal Cord".
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011810; BAA31514.1;
 DR InterPro: IPR003598; Ig_c2.

DR InterPro: IPR003600; Ig_1ike.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00408; IgC2; 2.
 DR SMART: SM00410; Ig_1ike; 1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 344 AA; 37613 MW; 22CAABF526A6B57E CRC64;
 Query Match 6.6%; Score 137.5; DB 13; Length 344;
 Best Local Similarity 26.0%; Pred. No. 0.0004;
 Matches 70; Conservative 38; Mismatches 104; Indels 57; Gaps 14;
 QY 3 RHLLTPEAVGSSGSGNEVI-EGPQNAVTLKSGQARFNCYSGQKLMALSLMVLVSR 53
 DB 20 RLFLVTPGVAVRSGDTPFKAMDNTVROGESATLRCSDVNRVAVLNRSLTYGN 79
 QY 54 DMVLSVRPEPIITNDRFTSQRYDQGNFTSEMIHNVESDSGNGINCSLO-----NS 107
 DB 80 DKWCIDPRVLLANTKTQYSIO-----IHVDVYDESPYCSQVQIDNHPTSVH-- 126
 QY 108 RLHGSAYLVQVNGELFIPSVNLVVAENPECEVTCLPSSHTRLPD--ISWELGLVSHS 165
 DB 127 RVH-----LIVQSPKITEISSISINEGNNISLFCIAIGRPE-PLTVWR--HISPA 176
 QY 166 YFVPEPSDQSAVSLATLPQSNGLTLCVATWKSILKARKSATVNLTVICP-----IRCPD 220
 DB 177 VGFISDEYLE-----ITGTIREQSGEYECAS-NDVAAPVQVRKVTAVNPPIFEAKGT 231
 QY 221 GGINIPGVL-----SLPSLGFSLPTWCK 245
 DB 232 GVPVGOKGILMCEASAVPSAEQO---WFK 257
 RESULT 8
 ID 09D4K2 PRELIMINARY; PRT; 173 AA.
 AC 09D4K2:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 4931420D14RIK protein.
 GN 4931420D14RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guinacch S., Hill D., Holman M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlski S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection".
 RL Nature 409:685-690(2001).
 DR EMBL; AK016467; BAB30253.1;
 DR MGD; MGI:1913992; 4931420D14RIK.
 SQ SEQUENCE 173 AA; 19581 MW; 147B6F155AC29FDF CRC64;

Query Match 6.6%; Score 136.5; DB 11; Length 173;
 Best Local Similarity 31.5%; Pred. No. 0.0002;
 Matches 40; Conservative 16; Mismatches 52; Indels 19; Gaps 6;

QY 260 CTTTIRCCCGRRCCGCGC-CGRCGCCRRRGRFIOFKSKSEKTKNETESGNS 318
 55 CSTRSCCCRCRC-CYCACCRC--CCSRFRFRSTTLRAVDPRFFQ--ITEKESL 109
 DB 319 GYNSDEQKTE-----TASLPKSCSSDPEQRNCGPPHARADQPPAPASHPQ 369
 110 GRRIRQLTFSQLELIEPEPTMALEPSEITVAFFSHKNANVSDP---EEVPECLDSDF 165
 QY 370 ASFWLAS 376
 DB 166 PNGDLAS 172

RESULT 9
 Q92626 PRELIMINARY; PRT; 1496 AA.

AC Q92626; PRELIMINARY; PRT; 1496 AA.
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE MYELOBLAST KIAA0230 (Fragment).
 CN KIAA0230.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=bone marrow;
 RX MEDLINE=97191544; PubMed=9039502;
 RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
 Tanaka A., Seki H., Miyajima N., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. VI.
 The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
 RT analysis of cDNA clones from cell line KG-1 and brain.";
 RL DNA Res. 3:321-329(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MELANOMA;
 RX MEDLINE=95048383; PubMed=7959781;
 RA Trent J.M., Taylor S.M., Deans R.J., Kan-Mitchell J., Mitchell M.S.,
 RT "Assignment of a human melanoma associated gene MG50 (D2S448) to
 RT chromosome 2p25.3 by fluorescence in situ hybridization.";
 RL Genomics 22:243-244(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MELANOMA;
 RA Mitchell M.S., Kan-Mitchell J., Miney B., Edman C., Deans R.J.;
 RT "Identification of a novel melanoma gene (MG50) - likely the gene for
 RT IL-1 receptor antagonist - which encodes epitopes recognized by human
 RT cytolytic T lymphocytes.";
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D86983; BAA13219.1; -
 DR EMBL: AF200348; AAF06354.1; -
 DR HSSP: P05184; ICKP.
 DR InterPro: IPR002007; Anim_peroxidase.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR003572; LRR_Nterm.
 DR InterPro: IPR003591; LRR_Typ.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF03098; An_peroxidase; 1.
 DR Pfam: PF00047; Ig; 4.
 DR Pfam: PF00560; LRR; 5.
 DR Pfam: PF01463; LRRCT; 1.

DR Pfam: PF00093; VWC; 1.
 DR PRINTS: PRO0457; ANPEROXIDASE.
 DR SMART: SM00408; IGC2; 4.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00369; LRR_Typ; 4.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWFC; UNKNOWN_1.
 KW Immunoglobulin domain.
 FT NON_TER 1
 SQ SEQUENCE 1496 AA; 167209 MW; E9B9A7069EFLABFF CRC64;

Query Match 6.5%; Score 135; DB 4; Length 1496;
 Best Local Similarity 25.0%; Pred. No. 0.0046;
 Matches 73; Conservative 36; Mismatches 127; Indels 56; Gaps 15;

QY 13 GSGSGNEVEFGPNATVVLKGSQARFCYTS-QGKMLIMWALSDMVVLSVRPEPIITDR 71
 DB 354 GSPARPTFVIQPNTEVLVESVTECSATGHPPRISMTRGDRFLPYDPVNTPTPS-- 411
 QY 72 FTSQRYDQGNFTSEMIINHPSPSGNIRCSLONS--RLHGSAYLVQVMGELFIPSVN 129
 DB 412 -----GG-----LYINVOGDSGEYACSAFNINIDSVAHTAFIYQALPQFTVTPQD 458
 QY 130 LVVAENPECEVTCLPSHMTLPDISW-ELGLVSHSYTFVEPESDLSAASITALLTPQS 188
 DB 459 RVYIEGQIVDFQC-EAKGNPPVIAWTKGSQLSVDRRLVLSGTLR--ISGVALLHDQ- 514
 QY 189 NGTLTGVATKSLKARKSATVNTLV-----IRCPQDT-----GGGINIPVLSLPSL 236
 DB 515 -GGYEQQAV--NITGSKVVAHLTVQPPVPAFASIPSTIYEVGANNVLP-----CSSQ 566

QY 237 GFSLP--TWGKVLGL-AGTMLLPPTCTLT-----RCCCRRCRC 275
 DB 567 GEPEPATWTKDGVQVTEGSKFHSPEGLTINDVPADAGRECVARNWIG 618

RESULT 10
 Q9P121 PRELIMINARY; PRT; 344 AA.

AC Q9P121;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Neurotrophin.
 GN NNT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Li G., Jin J., Tan X., Hu S., Yuan J., Qiang B.;
 RT "Cloning and identification of human neurotrophin full length cDNA.";
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF126426; AAR37591.1; -
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00408; IGC2; 2.
 DR SMART: SM00410; Ig_Like; 1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 344 AA; 37971 MW; DA4D12C295AABBE3A CRC64;

Query Match 6.4%; Score 133.5; DB 4; Length 344;
 Best Local Similarity 24.5%; Pred. No. 0.00092;
 Matches 65; Conservative 40; Mismatches 111; Indels 49; Gaps 13;

QY 3 RHLITVPAVSGSGNEVT-EGPNATVVLKGSQARFCYTSQSGKMLIMWALSDMVVLS-- 59
 DB 20 RLFLVPTGVAVRSGDATFPKAMDNVTVROGESATLCTIDNRVAVMLNRSTILYAGN 79

QY 60 -----VAPMEPIITNDREFTSORDGNGFTSEMIITHNVEPSDGNTRCSLQ-----NSRL 109
 Db 80 DKMCLDPRVVLN-----TQYX-----SLEIQNVAVYDEGPYCSVQJTNHPTKTSV 128
 QY 110 HGSAYLTVQVWGLFTIPSNVLYAENEPCEVCLPSHMTLRDLDISMELLOLASHSHSYFV 169
 Db 129 H-----LIVQVSKFVIVISSDISINEGNNISLFCIATGRPE-PLVYWR--HISPRAYGV 180
 QY 170 PEPDSLQSAVSILALTPQSNGLTLCVATWKSILKARKSATVNLTRCP-----ODTGGGI 224
 Db 181 SEDEYLE-----LOGITREGSGVECSAS-NDVAAPVVRKVTYVNPRTISAKGTGPV 235
 QY 225 NIPGVL-----SLPISLGSFLPTWGX 245
 Db 236 GQKGLQCEASAVPSAEQ---WYK 257

RESULT 11

QY 091562 PRELIMINARY; PRT; 1427 AA.
 AC 091562;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Tumor suppressor.
 GN XDCOA.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8335;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95113183; PubMed=7813784;
 RA Pieceni E.R.; Keale M.A.; Candia A.F.; Wright C.V.; Cho K.R.;
 RT "Expression of a homologue of the deleted in colorectal cancer (DCC)
 RT gene in the nervous system of developing Xenopus embryos."
 RL Dev. Biol. 166:654-665(1994).
 DR EMBL: U10986; AAA70168.1; -
 DR HSSP; P40189; 180U.
 DR InterPro: IPR003962; FNIII-repeat.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003958; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003005; Ig_MHC.
 DR Pfam; PF00041; fn3; 6.
 DR Pfam; PF00047; Ig; 4.
 DR PRINTS; PR00014; FNTPERIII.
 DR SMART; SM00060; FN3; 6.
 DR SMART; SM00408; IGC3; 3.
 DR SMART; SM00410; IG_Like; 2.
 KW Immunoglobulin domain; Repeat.
 SQ SEQUENCE 1427 AA; 156533 MW; 61FEA12C8A674972 CRC64;

Query Match 6.4%; Score 133; DB 13; Length 1427;
 Best Local Similarity 25.7%; Pred. No. 0.0065;
 Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 21 IEQONATVLKGSQARFRCYSGW--KLIMWALSDMVLVSVRMEPIITNDRETSQRYD 78
 Db 243 LQRSNVVAIEGDVAVECAVS-GYPTPIVMQGD-----EPVPIRTR---KYS 288
 QY 79 QGNGFTSEMIITHNVEPSDGNTRC--SLQNSRLHGSAYLTVQVWGLFTIPSNVLYAENE 136
 Db 289 VLGG--SNLISNVTDDACATVATYKNDNTSESDLTVMVPPQPLNHPANLYAESM 346
 QY 137 PCEVTCL---PSHMTLRPLDISW-ELGLVSHSHSYFVPEPSDQSAVSILALTPQSNGT 191
 Db 347 DIEFECAVSGKPS-----PLVYKMTKNGEVIVPSDYFQIVGSMNR-----ILGLVKSDGY 397
 QY 192 LTVCA-----TWKSLKARKSATVNLTVI-RCPODTGGGINIPGVLS 232

Db 398 YQCIENENAGNIQYVAQLIIPDPAVPSLIPSAKRDV-----VPLVYSS 442

RESULT 12

QY 09DBP0 PRELIMINARY; PRT; 697 AA.
 AC 09DBP0;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 21, Last annotation update)
 DE Solute carrier family 34 (sodium phosphate), member 2.
 GN SLC34A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LUNG;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J.; Shingawa A.; Shibata K.; Yoshino M.; Itoh M.; Ishii Y.;
 RA Arikawa T.; Hara A.; Fukunishi Y.; Kono H.; Adachi J.; Fukuda S.;
 RA Aizawa K.; Izawa M.; Nishi K.; Kiyosawa H.; Kondo S.; Yamanaoka I.;
 RA Saito T.; Okazaki Y.; Gojohori T.; Bono H.; Kasukawa T.; Saito R.;
 RA Kadota K.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;
 RA Fleischmann W.; Gaasterland T.; Giesi C.; King B.; Koehli H.;
 RA Kuehl P.; Lewis S.; Matsuo Y.; Nikaido I.; Pesole G.; Quackenbush J.;
 RA Schriml L.M.; Staudt F.; Suzuki R.; Tomita M.; Wagner L.; Washio T.;
 RA Sakai K.; Okido T.; Furuno M.; Aono H.; Balderelli R.; Barsh G.;
 RA Blake J.; Boffelli D.; Bojunga N.; Carninci P.; de Bonaldo M.F.;
 RA Brownstein M.J.; Bull C.; Fletcher C.; Fujita M.; Gariboldi M.;
 RA Gustincich S.; Hill D.; Hofmann M.; Hume D.A.; Kamiya M.; Lee N.H.;
 RA Lyons P.; Marchionni L.; Mashima J.; Mazzarelli J.; Mombaerts P.;
 RA Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto N.;
 RA Sasaki H.; Sato K.; Schoenbach C.; Seya T.; Shibata Y.; Storch K.-F.;
 RA Suzuki H.; Toyooka K.; Wang K.H.; Weitz C.; Whitlaker C.; Wilming L.;
 RA Wysshaw-Boris A.; Yoshida K.; Hasegawa Y.; Kawaji H.; Kohlsuki S.;
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL; AK004832; BAB23600.1; -
 DR MGD; MGI:1342284; Slc34a2.
 DR InterPro: IPR003841; Na_P1_cotrans.
 DR Pfam; PF02690; Na_P1_cotrans; 1.
 DR TIGRfams; TIGR01013; 2a58; 1.
 SQ SEQUENCE 697 AA; 76244 MW; 2A7B9384857EF16F CRC64;

Query Match 6.3%; Score 131; DB 11; Length 697;
 Best Local Similarity 20.3%; Pred. No. 0.0039;
 Matches 88; Conservative 53; Mismatches 124; Indels 168; Gaps 22;

QY 5 LITVPEAVSGSGNEVIEGPONATVLK-----GSOARFRCYVS 42
 Db 380 LIMIVKLDS-----VLRG-QVATVTKTLNDFEPFAMLTGYLALVAGAMTIVOSS 433
 QY 43 QGKLMALSDMVLVSVRMEPIITNDRETSQRYDGNFTSEMIITHNVEPSDGNTRC 102
 Db 434 SVFTSAMTPILIGIVISIRAYPLTGSNI-----GTTTALIALASGNT--LRS 483
 QY 103 SLQNSRLHGSAYLTVQVWGLFTIPSNVLYAENEPCEVCLPSHMTLRP-----DIS 154
 Db 484 SIQIALCH-----EFNINSGI-----LIMPPIPTRLPIRLAKGLNIS 522
 QY 155 -----WELGLVSHSHSYFVPEPSDQSAVSILALTPQSNGLTLCVATWKSILKARKSATV 210
 Db 523 AKYRW---FAVYLLIFFEFTVP-----LTVFGLS-----LAGPVLVGVGPITL 564
 QY 211 LVVI-----RCPODTGGGINIPGVL---SLPISLGSFLPTWGXGLAGLMILTP 258
 Db 565 LLLVLICLRLMIOFCRPR-----ILPLKLRDWNFLPLMHSKLPWDNV-----ISLAT 611
 QY 259 TCTLTTRCCCRRRRCG-----CNC-CCRCFCGCRKRGFRLOFQKSEKKTNETE 312

Db 612 TC-FORCCCCCGRVCCWCGCKCCRCSCCD-----GGEEREEQ----- 656
 QY 313 SCNENSGYNSDQKTTERTASLPPKSCSSDPEGRNSCGPPHORADQPPRASHPOASF 372
 Db 657 -----DIPVKSAGAFDNAMSKRC-----QDEGKGOV 683
 QY 373 NLASPEKVSNTTV 385
 Db 684 EYLSMKALSNNTTV 696

RESULT 13
 Q92290 PRELIMINARY; PRT: 697 AA.
 AC 092290;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Type IIB Na/phosphate-cotransporter.
 DE SLC34A2 OR NP12B.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NC NCBL_TaxID=10090;
 RX MEDLINE=99045724; PubMed=9826740;
 RA Hillebrand H., Hattenhauer O., Traebert M., Forster I., Murer H.,
 RA Biber J., Characterization of a murine type II sodium-phosphate cotransporter
 RT expressed in mammalian small intestine.
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14564-14569(1998).
 DR EMBL: AF081499; AAC80007.1;
 DR MGD: MGI:1342284; SLC34a2.
 DR InterPro: IPR003841; Na.Pi.cotrans.
 DR Pfam: PF02690; Na.Pi.cotrans. 1.
 DR TrEMBL: T1R01013; 2a58.1.
 DR TIGR: TIGR01013; 76286 MW: 83955CCE0F565265 CRC64;
 SQ SEQUENCE 697 AA; 76286 MW: 83955CCE0F565265 CRC64;

Query Match 6.38; Score 131; DB 11; Length 697;
 Best Local Similarity 20.38; Pred. No. 0.0039;
 Matches 88; Conservative 53; Mismatches 124; Indels 168; Gaps 22;

QY 5 LTUVEANGSGNVEITGPNATVTK-----GSOARFNCTVS 42
 Db 380 LTUVEANGSGNVEITGPNATVTK-----GSOARFNCTVS 43
 QY 43 OGKMLIMALSDMYLSVRPMEPIITNDRTSORYDOGNFTSEMIHNVSPDSGNIRC 102
 Db 434 SVFTSAMPTLIGIVISIERAVPLTGSNI-----GTTTALMALASPGNT--LRS 483
 QY 103 SLONSRHGSAYLVQVMGELFIPSVNLVAENECVETCLPSMTNRLP-----DIS 154
 Db 484 SLQIALCH-----FFENISGI-----LLWYPIPTRLPLRLANGLONIS 522
 QY 155 -----WELGLVHSSYFVEPESDLOSASIALTPQSNGLTCAVATWSLKARKSATVN 210
 Db 523 AKYRW---FAVRYLIFFEVTP-----LTVFGLS-----LAGPVLVGVGPITL 564
 QY 211 LTVI-----RCQDPTGGGINIPGV---SSLPISLGSFLPTWKGKGLAGTMLLTP 258
 Db 565 LLLVLCILMLNLFRCPR-----ILPLKLRDMNPLPLMWHSLKPDVNY-----ISLAT 611
 QY 259 TCTLTIRCCCCRRRCG-----CNC-CGRCCFCRRKRGFRIFQPKSEKKTNETETE 312
 Db 612 TC-FORCCCCCGRVCCWCGCKCCRCSCCD-----QDEGKGOV 683
 QY 313 SCNENSGYNSDQKTTERTASLPPKSCSSDPEGRNSCGPPHORADQPPRASHPOASF 372
 Db 657 -----DIPVKSAGAFDNAMSKRC-----QDEGKGOV 683
 QY 373 NLASPEKVSNTTV 385

Db 684 EYLSMKALSNNTTV 696
 : : : : :
 RESULT 14
 Q02048 PRELIMINARY; PRT: 164 AA.
 AC 022048;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE T01B7.8 protein.
 DE T01B7.8.
 GN Caenorhabditis elegans.
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCBL_TaxID=6239;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Sims M.A.;
 RA Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z66499; CA91301.1;
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR001271; Defensin_mammal.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001007; WVF_C.
 DR PROSITE: PS00158; 4Fe4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE: PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE: PS00202; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01208; WVF_C; UNKNOWN_1.
 DR TIGR: TIGR01013; 76286 MW: 83955CCE0F565265 CRC64;
 SQ SEQUENCE 164 AA; 16499 MW: 16499 MW: C002D4BD36C9FECED CRC64;

Query Match 6.28; Score 129.5; DB 5; Length 164;
 Best Local Similarity 32.18; Pred. No. 0.00081;
 Matches 43; Conservative 11; Mismatches 53; Indels 27; Gaps 6;

QY 157 LGLVHSSYFVEPESDLOSASIALTPQSNGLTCAVATWSLKARKSATVNLTVIRC 216
 Db 6 LGLVHSSYFVEPESDLOSASIALTPQSNGLTCAVATWSLKARKSATVNLTVIRC 216
 QY 217 PDGTGGGINIPGVLSLPSLGSFLPTWKGKGLAGTMLLPTTCLTIRCCCCRRRCGG 276
 Db 58 KRGGCGCGCGCGC-----GCCGCGGCGGCG--CGCCCRPRCCCGCCRCCTC 101
 QY 277 --NCCC-RCFCRCR 287
 Db 102 CRTCCCTCCTCCR 115

RESULT 15
 Q01761 PRELIMINARY; PRT: 6632 AA.
 AC 001761;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE C. elegans UNC-89 protein (corresponding sequence C09D1.1).
 DE UNC-89.
 GN Caenorhabditis elegans.
 OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCBL_TaxID=6239;
 RX MEDLINE=99069613; PubMed=9851916;
 RA STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;

RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du Z., Le T.T., Wilson R.;
 RT "The sequence of C. elegans cosmid C09D1.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003131; AAB54132.2;
 SO SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960E89 CRC64;

Query Match
 Best Local Similarity 23.0%; Score 129; DB 5; Length 6632;
 Matches 59; Conservative 41; Mismatches 97; Indels 60; Gaps 10;

OY 6 LTVPEAVGSG-----SGNE-----VIEGPQNAVYLGSGQARFNC 40
 DB 2038 LIPMAQDSKITEVENSSESSNQLTVPNPSITPIVVDGPKSVTIKETETAEFKAT 2097
 OY 41 VSQGM--KLIMWALSDMAYLVSRPMEPIITNDRFTSORYDGGNFTSEMIHNVEPSDSG 98
 DB 2098 IS-GFPAPTYKWTINEKIVESTRITITIKTEDVYT-----LKISMAKIEQTC 2143
 OY 99 NIRCSONSRILHGSAYLVQVMEGLFIPS-----VNIYVAENRCEVTC--PSHWRL 150
 DB 2144 TVKVTAGNSAGQSKQADLVEPNVAKPKFSQLTIDKVADEGEPLRWNLLELDGPPSGT-- 2201
 OY 151 PDISMEL-GLIVSHSYFYVEPSDLOSASVILALTPQSNGLTVCVATWKSILKARKSATV 209
 DB 2202 -EVSWMINGQPLKSDTVQVVDHGDGYHTVIAEKPEMSGTLLAKAKNAAGECETSAKV 2260
 OY 210 NLT-----VIRCPD 219
 DB 2261 TVNGNKKPEFVQAPON 2277

Search completed: April 28, 2003, 21:07:42
 Job time : 41.1283 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 28, 2003, 16:24:30 ; Search time 10.3692 Seconds

(without alignments)
1543.990 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077
Sequence: 1 MERHLITVPAVSGSGSNGEV.....HPQASFNLASPEKSVNTTVV 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	7.3	404	1 RAGE_HUMAN	Q15109 homo sapien
2	139.5	6.7	344	1 NTRI_RAT	Q62718 rattus norv
3	133	6.4	337	1 G55A_CHICK	Q98892 gallus gall
4	131	6.3	1447	1 DCC_MOUSE	P70211 mus musculu
5	130	6.3	345	1 OPCM_HUMAN	P11834 bos taurus
6	129	6.2	345	1 OPCM_BOVIN	Q90773 gallus gall
7	129	6.2	353	1 CEPV_CHICK	P32736 rattus norv
8	128	6.2	345	1 OPCM_RAT	F43146 homo sapien
9	128	6.2	1447	1 DCC_HUMAN	Q15746 homo sapien
10	128	6.2	1914	1 KML5_HUMAN	P98160 homo sapien
11	125	6.0	1070	1 PTK7_HUMAN	Q13308 homo sapien
12	122.5	5.9	4393	1 PGBM_HUMAN	P92154 coturnix co
13	118.5	5.7	620	1 SMP_COTUJA	Q90610 gallus gall
14	117.5	5.7	1443	1 NEOL_CHICK	Q92859 homo sapien
15	114.5	5.5	1461	1 NEOL_HUMAN	Q28173 bos taurus
16	113.5	5.5	416	1 RAGE_BOVIN	P97603 rattus norv
17	113.5	5.5	1377	1 NEOL_RAT	P78310 homo sapien
18	111	5.3	365	1 CXAR_HUMAN	Q62556 mus musculu
19	110.5	5.3	524	1 BUTY_MOUSE	Q08180 drosophila
20	110.5	5.3	764	1 ICCR_DROME	Q03751 drosophila
21	109.5	5.3	249	1 CSP_DROME	Q06418 homo sapien
22	109.5	5.3	880	1 TYO3_HUMAN	P36335 xenopus lae
23	106	5.1	1092	1 NCAL_XENLA	P05793 mus musculu
24	105.5	5.1	862	1 CD22_MOUSE	Q05793 mus musculu
25	105.5	5.1	3707	1 PGBM_MOUSE	Q99795 homo sapien
26	105	5.1	319	1 A33_HUMAN	P15664 drosophila
27	104.5	5.0	333	1 AMAL_DROME	Q62786 rattus norv
28	104.5	5.0	879	1 FPRP_RAT	Q98919 gallus gall
29	103.5	5.0	338	1 LAMP_CHICK	P08921 rattus norv
30	102.5	4.9	344	1 CD2_RAT	Q92018 rattus norv
31	102.5	4.9	348	1 KILO_RAT	P97792 mus musculu
32	102.5	4.9	365	1 CXAR_MOUSE	Q14162 homo sapien
33	100.5	4.8	830	1 SRBC_HUMAN	

ALIGNMENTS

RESULT 1	RAGE_HUMAN	STANDARD	PRT	404 AA.	
AC	Q15109: Q15279; Q9Y3R3; Q9H2X7;				P11799 gallus gall
DT	01-NOV-1997 (Rel. 35, Created)				Q06561 caenorhabdi
DT	01-NOV-1997 (Rel. 35, Last sequence update)				P34082 drosophila
DT	15-JUN-2002 (Rel. 41, Last annotation update)				P14781 gallus gall
DE	Advanced glycosylation end product-specific receptor precursor				Q60500 homo sapien
DE	(Receptor for advanced glycosylation end products).				P55146 rattus norv
GN	AGER OR RAGE.				P11627 mus musculu
OS	Homo sapiens (Human).				Q13449 homo sapien
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Q62813 rattus norv
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				Q24372 drosophila
OX	NCBI_Taxid=9606;				P08169 bos taurus
RN	[1]				P13596 rattus norv
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Lung;				
RX	MEDLINE=92340547; PubMed=1378843;				
RA	Neepser M., Schmidt A.M., Brett J., Van S.D., Wang F., Pan Y.C.,				
RA	Elliston K., Stern D., Shaw A.;				
RT	"Cloning and expression of a cell surface receptor for advanced				
RT	glycosylation end products of proteins.";				
RL	J Biol. Chem. 267:14998-15004(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=95137587; PubMed=7835890;				
RA	Sugaya K., Fukagawa T., Matsumoto K., Mila K., Takahashi E., Ando A.,				
RA	Inoko H., Ikemura T.;				
RT	"Three genes in the human MHC class III region near the junction with				
RT	the class II: gene for receptor of advanced glycosylation end				
RT	products, PBX2 homeobox gene and a notch homolog, human counterpart				
RL	of mouse mammary tumor gene int-3.";				
RN	Genomics 23:408-419(1994).				
RP	[3]				
RA	SEQUENCE FROM N.A. (ISOFORM 1).				
RA	Rosen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,				
RA	Banta A., Spies T., Hood L.;				
RT	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.				
RL	[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.				
RA	Abelton M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,				
RA	Yamamoto H.;				
RT	"Molecular heterogeneity of the receptor for advanced glycation				
RT	endproducts.";				
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RP	[5]				
RA	SEQUENCE FROM N.A. (ISOFORM 2).				
RA	Malherbe P., Richards J., Gailiard H., Thompson A., Diener C.,				
RA	Schlier A., Huber G.;				
RT	"cDNA cloning of a novel secreted isoform of the human Receptor for				
RT	Advanced Glycation End products (RAGE) and characterization of cells				
RT	co-expressing cell-surface scavenger receptors and Swedish mutant				
RL	amyloid precursor protein.";				
RP	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				

RC TISSUE=Lung;
 RA Strausberg R.;
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE OF 1-12 FROM N.A.
 RA Hudson B.I., Futers T.S.;
 RT "Novel polymorphisms in the receptor for advanced glycation
 end-products (RAGE) gene";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATED PROTEINS
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 CC RATE IN DIABETES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
 CC Secreted (isoform 2).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2/RAGESEC;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M91211; AAA03574.1; -;
 CC EMBL; D28769; BAA05958.1; -;
 CC EMBL; U89336; AAB47491.1; -;
 CC EMBL; AB036432; BAA89369.1; -;
 CC EMBL; AJ133822; CAB43108.1; -;
 CC EMBL; BC020669; AAB20669.1; -;
 CC EMBL; AF208289; AAG55728.1; -;
 CC Genew; HGNC:320; AGER.
 CC MIM; 600214; -;
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003598; Ig_C2.
 CC InterPro; IPR003600; Ig_Like.
 CC Pfam; PF00047; Ig_2.
 CC SMART; SM00410; Ig_Like; 1.
 CC SMART; SM00408; IGC2; 1.
 CC PROSITE; PS00290; IG_MHC; 1.
 CC Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
 CC Alternative splicing; Polymorphism.
 CC SIGNAL 1 22 POTENTIAL.
 CC CHAIN 23 404 ADVANCED GLYCOSYLATION END PRODUCT-
 CC SPECIFIC RECEPTOR.
 CC DOMAIN 23 342 EXTRACELLULAR (POTENTIAL).
 CC TRA5MEM 343 363 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 364 404 IG-LIKE V-TYPE DOMAIN.
 CC DOMAIN 31 106 IG-LIKE C2-TYPE DOMAIN 1.
 CC DOMAIN 137 215 IG-LIKE C2-TYPE DOMAIN 2.
 CC DOMAIN 252 308 POTENTIAL.
 CC DISULFID 38 99 POTENTIAL.
 CC DISULFID 144 208 POTENTIAL.
 CC DISULFID 259 301 POTENTIAL.
 CC CAROHD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CAROHD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CAROHD 81 81 POLY-GLU.
 CC VARSPIC 54 67 MISSING (IN ISOFORM 2).
 CC VARSPIC 275 404 GVPLPSPVLLPEIGPODGYSCVATSHSGPOESRA
 CC VARSPIC 275 404 VSIITGEGSPAGSGGSLGTLALGIGLGTAL
 CC LIGVILMORRORRERKAPRNOEERERATLNSEPEAC
 CC ESSYGGP -> VSDLERGAGTRRGAGRCILGRIKGNSS
 CC PGPGDPRGDSRAHMGHLVAKAATPRGEGRRKGGRG
 CC GACREYVGGT (IN ISOFORM 2).
 CC O -> R.
 CC /FTID=VAR_011338.
 CC M -> G (IN REF. 1).
 CC CONFLICT 1 1

SQ SEQUENCE 404 AA; 42802 MW; 0D584C436C30CCE7 CRC64;
 Query Match 7.3%; Score 151; DB 1; Length 404;
 Best Local Similarity 23.3%; Pred. No. 7e-05;
 Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19;
 QY 9 PEAVSGSGNEVIEGPONATVTKSGQARNCTVSG--WKILMALSDMYLVSRNEP 65
 DB 124 PEIVDSAS--ELTAGVPN-----KVTGVSSEGSYPAGTILSWHLDG-- 163
 QY 66 IITNDFTS-----QRYDDGNGFT--SEMIINHVPSGNGIR-----CSLQSRHLGSAY 114
 DB 164 LVNNEKGVSVKEQTRRHPTGLFTLQSELM--VTPARQGDPRPTFSCSFSGLPRIHAL 220
 QY 115 LTVQVAGLELIP--SVNLVVAENP-----CEVTLPSHMTPLDISE 156
 DB 221 RTAPIDPRWEPVPLEEVQV---EPGGAVALPGGTVTLTCEVPAQS---PQIHMM 272
 QY 157 LGLVSHSSYTFPPERSDQASVIALTPQSNGLTICVATWKSILKARSATVNLTVIRC 216
 DB 273 KD-----GVPLPSPVLLPEIGPODGYSCVATSHSGPOESRAVSISIE- 322
 QY 217 PODTGGGINIPVLSLPSLGFSLPTWGRVGLIAGTMLLT-----PTCTLTIRCCC 269
 DB 323 PGEEG-----PTAGSVGSGGLGTALALGILGGLTALLGVLIM 363
 QY 270 RRRCCGCCNCCRCFCRCRRKGRFIOFKKSEKKT--NKEPETSNGNSGYNSDOKT 327
 DB 364 QRR-----QRRGERKAPNEQEEERALELN----- 389
 QY 328 TETASLPKSCSSDPEQRNNSGCP 353
 DB 390 -----QSEEPAGESSYGGP 404
 RESULT 2
 NTRL_RAT
 ID NTRL_RAT STANDARD; PRT; 344 AA.
 AC 062718;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Neurotrophin precursor (GP65).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.
 RC SPRAIN-Sprague-Dawley;
 RX MEDLINE=95198094; PubMed=7891157;
 RA Struyk A.F., Canoll P.D., Wolfigang M.J., Rosen C.L., D'Eustachio P.,
 RT Salzer J.L.;
 RT "Cloning of neurotrophin defines a new subfamily of differentially
 expressed neural cell adhesion molecules.";
 RL J. Neurosci. 15:2141-2156 (1995).
 CC -1- FUNCTION: NEURAL CELL ADHESION MOLECULE.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN SEVERAL
 CC DEVELOPING PROJECTION SYSTEMS: IN NEURONS OF THE THALAMUS,
 CC SUBPATE, AND LOWER CORTICAL LAMINAE IN THE FOREBRAIN AND IN THE
 CC PONTINE NUCLEUS, CEREBELLAR GRANULE CELLS, AND PURKINJE CELLS IN
 CC THE HINDBRAIN.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. (GLON
 CC SUBFAMILY).
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U16845; AAA67445.1; -

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_Like.

DR Pfam: PF00047; Ig; 3.

DR SMART: SM00410; Ig_Like; 1.

DR SMART: SM00408; IgC2; 2.

KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;

KW Repeat: Signal.

FT SIGNAL 1 31

FT CHAIN 32 321

FT PROPEP 322 344

FT DOMAIN 50 122

FT DOMAIN 150 208

FT DOMAIN 236 302

FT DISULFID 57 115

FT DISULFID 157 201

FT DISULFID 243 295

FT CARBOHYD 44 44

FT CARBOHYD 70 70

FT CARBOHYD 152 152

FT CARBOHYD 216 216

FT CARBOHYD 284 284

FT CARBOHYD 292 292

FT CARBOHYD 305 305

FT CARBOHYD 321 321

FT CARBOHYD 321 321

SO SEQUENCE 344 AA; 37998 MW; CBB39BE53B3B224 CRC64;

Query Match

Best Local Similarity 24.9%; Score 139.5; DB 1; Length 344;

Matches 66; Conservative 40; Mismatches 110; Indels 49; Gaps 13;

QY 3 RHILTVENAVGSSGNGEYI-EGPONATVYLGSGARFNCYSGKGLIMMALSDMYVL-- 59

DB 20 RLFLVLPVGVVRSRGGATFPFAMDNVTVRGESATLRCITDNRYAVAMNRSITLACN 79

QY 60 ---VRPEPIITNDRTSORYDQGNFTSEMIHNVEPDSGNGIRCSLQ-----NSRL 109

DB 80 DKWCLDRVYVLNS---TQYQY-----SIEIQNDVYDEGPHYTCSVTIDNHPTSKRV 128

QY 110 HGSAYLVQVWGEFLFIPSVNLVVAENPECEVYCLPSHWTRLPDISWELGLVSHSYFEV 169

DB 129 H----LIVQVSPKIVEISDISISINEGNINISLTCIATGRPE-PTVYWR---HISPKAVGFV 180

QY 170 PEPDLSQAVSIALTPQSNGLTLCVATWKSLSKARKSATVNLVYICRP-----QDTGGGT 224

DB 181 SEDEYLE---TQGITREQSGEYECAS-NDVAPVPRVAVNVVPPYISEAKGTGVPV 235

QY 225 NIPGVL---SSLPGLGFSLPYTWK 245

DB 236 GQKGTLCCEASAVPSAEFG---WFK 257

RESULT 3

G55A_CHICK

ID G55A_CHICK STANDARD; PRT; 337 AA.

AC 098892;

DT 01-NOV-1997 (Rel. 35, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Neurite inhibitor GP55-A precursor (OBSCAM protein gamma isoform).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Lodge A.P., Kim D.-S., Howard M.R., McNamee C.J., Smith N., Moss D.J.;

RT "Cloning of CEPU-9, a secreted isoform of CEPU-1, and OBSCAM CDNA's from

RT chick: structural diversity of Iglon family proteins.";

RL Submitted (Apr-1999) to the EMBL/Genbank/DBD databases.

RN [2]

RP SEQUENCE OF 64-337 FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Brain;

RX MEDLINE-97157768; PubMed-9004047;

RA Wilson D.J.A., Kim D.-S., Clarke G.A., Marshall-Clarke S., Moss D.J.;

RT A family of glycoproteins (GP55), which inhibit neurite outgrowth,

RT are members of the Ig superfamily and are related to OBSCAM,

RT neurotrophin, IAMP and CEPU-1.";

RL J. Cell Sci. 109:3129-3138(1996).

CC -1- FUNCTION: INHIBITS NEURITE OUTGROWTH.

CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE NERVOUS SYSTEM.

CC -1- DEVELOPMENTAL STAGE: INCREASES DURING DEVELOPMENT FROM VERY LOW

CC LEVELS AT EMBRYONIC DAY 10 AND IS MOST ABUNDANT AFTER HATCHING.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON

CC SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: Y08170; CAB41420.1; -

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_Like.

DR Pfam: PF00047; Ig; 3.

DR SMART: SM00410; Ig_Like; 1.

DR SMART: SM00408; IgC2; 2.

KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;

KW Repeat: Signal.

FT SIGNAL 1 20

FT CHAIN 21 317

FT PROPEP 318 337

FT DOMAIN 43 115

FT DOMAIN 143 201

FT DOMAIN 229 295

FT DISULFID 50 108

FT DISULFID 150 194

FT DISULFID 236 288

FT CARBOHYD 133 133

FT CARBOHYD 277 277

FT CARBOHYD 285 285

FT CARBOHYD 298 298

SO SEQUENCE 337 AA; 36887 MW; BAE717551856551E CRC64;

Query Match

Best Local Similarity 6.4%; Score 133; DB 1; Length 337;

Matches 71; Conservative 36; Mismatches 107; Indels 60; Gaps 16;

QY 5 LITVPEAVGSSGNGEYI-EGPONATVYLGSGARFNCYSGKGLIMMALSDMYVL----- 58

DB 16 LFLFIP-GVPRSGDATFPFAMDNVTVRGESATLRCITDNRVRAVA-LNRSITLYAGND 73

QY 59 ---SVRPEPIITNDRTSORYDQGNFTSEMIHNVEPDSGNGIRCSLQ-----NSRLH 110

DB 74 KMSIDRRVYVLNS---TKQY-----SIEIHNVDYDEGPHYTCSVTIDNHPTSKRVH 122

QY 111 GSAIVLVQVWGEFLFIPSVNLVVAENPECEVYCL---PSHWTRLPDISWELGLVSHSY 166

DB 123 ---LIVQVSPKIVEISDISITVNEGSSVTLKCIATGRPE---PTVYWR---HISGKGQ 170

QY 167 YFVPEPDLQAVSIALTPQSNGLTLCVATWKSLSKARKSATVNLVYICRP-----IRCPDGT 221

DB 171 GFVSEDEYLE---ITGITREQSGEYECASV-NDVAPVPRVAVNVVPPYISEAKGTGVPV 225

OY 222 GGINIPGVL-----SSLPISGLFSLPTGKVGGLA 251
 DB 226 ASVGKGKILQCEASAVPAEFO---WKEEDIRLA 256

RESULT 4
 DCC_MOUSE STANDARD: PRT: 1447 AA.
 ID DCC_MOUSE P70211:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor suppressor protein DCC precursor.
 GN DCC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_Taxid=10090;
 RX SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 RC STRAIN-BALB/c; TISSUE-Brain;
 RA MEDLINE=96112625; PubMed=9570174;
 RA Cooper H.M., Armes P., Britto J., Gad J., Wilks A.F.;
 RT "Cloning of the mouse homologue of the deleted in colorectal cancer
 RT gene (mdcc) and its expression in the developing mouse embryo.";
 RT Oncogene 11:2243-2254(1995).
 RL [2]
 RP REVISIONS.
 RC STRAIN-BALB/c; TISSUE-Brain;
 RA Cooper H.M.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: A (shown here) and B; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
 CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
 CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
 CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
 CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION
 CC AND REMAIN AT THIS LEVEL IN THE ADULT.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X85788; CAA59786.1; -
 CC HSSP: P56276; 1TKL.
 CC MGD: MGI:94869; DCC.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003006; 19_MHC.
 DR InterPro: IPR003598; 19_c2.
 DR InterPro: IPR003600; 19_Like.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; 19_4.
 DR PRINTS: PR00014; ENTPEPITI.
 DR SMART: SM00060; FN3; 6.
 DR SMART: SM00410; IG_Like; 2.
 DR SMART: SM00408; IGc2; 3.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Repeat; Anti-oncogene; Alternative Initiation; Alternative splicing.

FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT CHAIN 85 1447
 FT INIT MET 85 85
 FT DOMAIN 26 1097
 FT TRANSMEM 1098 1122
 FT DOMAIN 1123 1447
 FT DOMAIN 54 124
 FT DOMAIN 154 219
 FT DOMAIN 254 317
 FT DOMAIN 345 407
 FT DOMAIN 426 522
 FT DOMAIN 525 618
 FT DOMAIN 619 716
 FT DOMAIN 722 816
 FT DOMAIN 840 940
 FT DOMAIN 941 1042
 FT DISULFID 61 117
 FT DISULFID 161 212
 FT DISULFID 261 310
 FT DISULFID 352 400
 FT CARBOHYD 60 94
 FT CARBOHYD 94 299
 FT CARBOHYD 299 318
 FT CARBOHYD 318 478
 FT CARBOHYD 478 628
 FT CARBOHYD 702 702
 FT VARSPLIC 819 838
 SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C22D5B9F CRC64;

POTENTIAL.
 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM A.
 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM B.
 FOR ISOFORM B.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE C2-TYPE DOMAIN 1.
 IG-LIKE C2-TYPE DOMAIN 2.
 IG-LIKE C2-TYPE DOMAIN 3.
 IG-LIKE C2-TYPE DOMAIN 4.
 FIBRONECTIN TYPE-III 1.
 FIBRONECTIN TYPE-III 2.
 FIBRONECTIN TYPE-III 3.
 FIBRONECTIN TYPE-III 4.
 FIBRONECTIN TYPE-III 5.
 FIBRONECTIN TYPE-III 6.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 MISSING (IN ISOFORM C).
 Query Match 6.3%; Score 131; DB 1; Length 1447;
 Best Local Similarity 25.2%; Pred. No. 0.014;
 Matches 55; Conservative 33; Mismatches 90; Indels 40; Gaps 9;

OY 9 PEAVGSGGNEV-----IEGQNFVTKGSGARNTVSGW--KLIMNAL 52
 DB 217 PASRTGNEAEVRRLSDPGHRLQYFLQRPSTVIAIEGDAVIECCVS-GYPPSTWLR 275
 OY 53 SDVVLSVRMEPIITNDRTSQRYPDGGNFTSEMIHNVPSDSGNIRC--SLQNSRLH 110
 DB 276 GEEVI-----QLRSKRYSLIG--SNLLISNVTDDSDGTVCVYVKNENIS 320
 OY 111 GSATVTVQVWGELFIPSVNLVAENECVETCPHSMTRLPDISW-ELGLIVSHSSYIV 169
 DB 321 ASAEITVLPVPWFLPHNSNYAYESMDIEFECAVS-GKPPVTVMMKNDVYIPSDYEI 379
 OY 170 PEPDLSAVSILALTPQNSGTLTCVATWKSARKSA 207
 DB 380 VGSNLR-----ILGVKSDGCFYQCAVNAENAGNAQSSA 413

RESULT 5
 OPM_HUMAN STANDARD: PRT: 345 AA.
 ID OPM_HUMAN Q14982;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OBICAM)
 DE Opioid-binding cell adhesion molecule (OPICML).
 GN OBICAM OR OBICAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_Taxid=9606;
 RX SEQUENCE FROM N.A.
 RC TISSUE-Occipital cortex;
 RX MEDLINE=95237612; PubMed=7721093;
 RA Shark K.B., Lee N.M.;
 RT "Cloning, sequencing and localization to chromosome 11 of a cDNA
 RT encoding a human opioid-binding cell adhesion molecule (OBICAM).";

RL Gene 155:213-217(1995).
 CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY
 CC INVOLVED IN CELL CONTACT.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J34774; AAA36387.1; -
 DR Genbank: HGNC:8143; OPMCML.
 DR MIM: 600632; -
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_3
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; Igc2; 2.
 DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 KW Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULEFID 57 115
 FT DISULEFID 157 202
 FT DISULEFID 244 296
 FT CARBOHYD 44 44
 FT CARBOHYD 70 70
 FT CARBOHYD 140 140
 FT CARBOHYD 285 285
 FT CARBOHYD 293 293
 FT CARBOHYD 306 306
 FT LIPID 322 322
 SQ SEQUENCE 345 AA; 38007 MW; E7AD17BEA1AA3FF4 CRC64;
 Query Match 6.3%; Score 130; DB 1; Length 345;
 Best Local Similarity 25.6%; Pred. No. 0.0028;
 Matches 73; Conservative 34; Mismatches 108; Indels 70; Gaps 17;
 OY 3 RHLLTPEAVGSGSGNEVI-EGPQNAIVLKSGQARFNCVSGWKLIMWALSDNVVL---58
 DB 20 RLLEFVPEVGVPRSGDAFPRKMDNVVIRQGSATLRCTIDRVRYVAM-LNRSTILLAG 78
 OY 59 ----SVRPEPELITNDRTSQRYDQGNFTSMITHVPEPSDGNIRSLQ-----NSR 108
 DB 79 NDKWSIDPRVITLVN---TPROY-----SIMIONVDYDEGPTYTSVGTDNHPTKRS 127
 OY 109 LHGSAVLTVOVWGELEFIPSVNLVVAENEPCEVTCL---PSHWRLPDISWELGLVSHS 164
 DB 128 VH---LIVQVAPQIMNISSDITVNEGSSVLLCLAIORPE-----PIVYMR---HL 172
 OY 165 SYV---FVPEPSDQSAVSIALLTPQSGNGLTVCAVTKSKARKSATVNLTV-----IR 215
 DB 173 SVKEGQGVSEDEYLE---ISDIKRDQSGEYCSAL-NDVAADVVRKVKITVAVPYIS 227
 OY 216 CPQDTGGGINIPGVL-----SSLPISLGFSLPTMGK-----VGGIAG 252
 DB 228 KAKNIGVGVGQKGIILSCASAVPMAEFO---WKEETRLATGLDG 269

OPCM_BOVIN
 ID OPCM_BOVIN STANDARD; PRT; 345 AA.
 AC P11834;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OBAM)
 DE (Opioid binding cell adhesion molecule) (OPCML).
 GN OPCML OR OBAM OR OCAM.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NX NCBI_TaxID=9913;
 RP [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=69251576; PubMed=2721489;
 RA Schofield P.R., McFarland K.C., Hayflick J.S., Wilcox J.N.,
 RA Cho T.M., Roy S., Lee N.M., Loh H.H., Seeburg P.H.;
 RT "Molecular characterization of a new immunoglobulin superfamily
 RT protein with potential roles in opioid binding and cell contact".
 RL EMBO J. 8:489-495(1989).
 CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY
 CC INVOLVED IN CELL CONTACT.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X12672; CAA31192.1; -
 DR PIR: S03199; S03199.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; Igc2; 2.
 DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 KW Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULEFID 57 115
 FT DISULEFID 157 202
 FT DISULEFID 244 296
 FT CARBOHYD 44 44
 FT CARBOHYD 70 70
 FT CARBOHYD 140 140
 FT CARBOHYD 285 285
 FT CARBOHYD 293 293
 FT CARBOHYD 306 306
 FT LIPID 322 322
 SQ SEQUENCE 345 AA; 37914 MW; D1ECB09E7D8CB19 CRC64;
 Query Match 6.2%; Score 129; DB 1; Length 345;
 Best Local Similarity 25.7%; Pred. No. 0.0034;
 Matches 72; Conservative 33; Mismatches 109; Indels 66; Gaps 16;
 OY 3 RHLLTPEAVGSGSGNEVI-EGPQNAIVLKSGQARFNCVSGWKLIMWALSDNVVL---58

Db 20 RLFLVPTGVYVRSQDGFPAKMDNTVYRQESATLCTIDDRTRAM-LINSTIIYAG 78
 QY 59 -----SVRPMEDITNDREFTSQRDQGNFPTSEMITIHVNEPDSGNIRCSIQ-----NSR 108
 Db 79 NDKMSIDPRVILIVN-----TPTQY-----SIMQNVVDYDEGPYICSVQTDNHPKTR 127
 QY 109 LHGSAIYLVQVWGLFIPSVNLVVAENEPCEVYTL-----PSHWRLDISELGLVSHS 164
 Db 128 VH-----LIVQVPPQIMNITSDVFNESGVTLCLAIGRPE-----PVTWR-----HL 172
 QY 165 SYV-----FVPPSDIOSAVSITALTPOSNGLTCVATWKSILKARKSATVNLTV-----IR 215
 Db 173 SYKEGQGVSEDEYLE-----ISLDIKRQSGEYECSSAL-NDVAADPVKVKITVWPPYIS 227
 QY 216 CPDGTGGGINIPGVL-----SSLPSLGFSLPTWKGVLGLA 251
 Db 228 KAKNTGVSVGQKGLISCEASAVPMAEFQ---WFKEDTRLA 264

RESULT 7
 CEPU_CHICK STANDARD; PRT; 353 AA.
 AC 090773;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CEPU-1 protein precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96370549; PubMed=8774445;
 RA Spaltmann F., Brumentdorf T.;
 RT "CEPU-1, a novel immunoglobulin superfamily molecule, is expressed by
 developing cerebellar Purkinje cells.";
 RL J. Neurosci. 16:1170-1179(1996).
 CC -1- FUNCTION: IT MAY BE A CELLULAR ADDRESS MOLECULE SPECIFIC TO
 CC PURKINJE CELLS. IT MAY REPRESENT A RECEPTOR OR A SUBUNIT OF A
 CC RECEPTOR COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A MAJOR FORM AND A MINOR FORM.
 CC (SHOWN HERE) ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: FOUND ON THE DENDRITES, SOMATA AND AXONS OF
 CC DEVELOPING PURKINJE CELLS. UNDETECTABLE ON OTHER NEURONS LIKE
 CC GOLGI OR GRANULE CELLS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BY DEVELOPING CEREBELLAR PURKINJE
 CC CELLS. EXPRESSION COINCIDES WITH THE GROWTH OF THE DENDRITIC TREE,
 CC AFTER PURKINJE CELLS HAVE FINISHED THEIR MIGRATION FROM THE
 CC VENTRICULAR ZONE (FROM E15 UNTIL E21). EXPRESSED IN THE ADULT.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: 272497; CA96578.1; -
 DR InterPro: IPR003006; IG_LMC.
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003600; IG_Like.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; IGC2; 2.

KW Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
 KW Repeat; Signal; Alternative splicing.
 FT CHAIN 29 330
 FT PROPE 331 353
 FT DOMAIN 48 120
 FT DOMAIN 148 206
 FT DOMAIN 234 300
 FT DISULFID 55 113
 FT DISULFID 155 199
 FT DISULFID 241 293
 FT CARBOHYD 42 42
 FT CARBOHYD 68 68
 FT CARBOHYD 150 150
 FT CARBOHYD 282 282
 FT CARBOHYD 290 290
 FT CARBOHYD 303 303
 FT CARBOHYD 330 330
 FT LIPID 310 320
 FT VARSPLIC 353 AA; 38736 MW; 2550CA8591EBBBA6 CRC64;
 SQ SEQUENCE

Query Match 6.2%; Score 129; DB 1; Length 353;
 Best Local Similarity 25.0%; Pred No. 0.0035;
 Matches 66; Conservative 38; Mismatches 92; Indels 68; Gaps 14;
 Db 35 TTPKAM-----DNVTVRQGESATLRCSVDNRYTRVAMLRRSSILYAGNKKWL 82
 QY 59 SVRPMEDITNDREFTSQRDQGNFPTSEMITIHVNEPDSGNIRCSIQ-----NSRLHGS 112
 Db 83 DPRVLIANTKTYQTSIQ-----IHVDYIDSGPYTCSVQTDNHPKTRSVH-- 127
 QY 113 AYLVQVWGLFIPSVNLVVAENEPCEVYTLPSHWRLPD--ISWELGLVSHSYFVP 170
 Db 128 --LIVQVSPKITESSDISINEGNVSLCIA--TGRDPTITWR--HISPKAVGFIS 179
 QY 171 EPSDLOSASVITALTPOSNGLTCVATWKSILKARKSATVNLTV-----TRCPDGTGGGIN 225
 Db 180 EDEYLE-----ITGITRQSGEYECSSAS-NDVAAPVQVRKVKVYVPPYISDAKSTGVPG 234
 QY 226 IPGVL-----SSLPSLGFSLPTWKG 245
 Db 235 QKGLMCEASAVPSADPQ---WYK 255
 RESULT 8
 OPCM_RAT STANDARD; PRT; 345 AA.
 ID OPCM_RAT
 AC P32736; 001654; P32735; 001653;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule precursor (OBPAM)
 DE (Opioid-binding cell adhesion molecule) (OPCML).
 OS OPCML OR OBPAM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92347701; PubMed=1339369;
 RA Lippman D.A., Lee N.W., Loh H.H.;
 RT "Opioid-binding cell adhesion molecule (OBPAM)-related clones from a
 RT rat brain cDNA library."
 RL Gene 117:249-254(1992).
 RN [2]
 RP SEQUENCE OF 195-214, AND GPI-ANCHOR.
 RX MEDLINE=95198094; PubMed=7891157;
 RA Struys A.F., Canoll P.D., Wolfigang M.J., Rosen C.L., D'Eustachio P.,
 RA Salzer J.L.;

RT Cloning of neurotrophin defines a new subfamily of differentially
 RT expressed neural cell adhesion molecules.*;
 RL J. Neurosci. 15:2141-2156(1995).
 CC -1- FUNCTION: BINDS OPIOIDS IN THE PRESENCE OF ACIDIC LIPIDS; PROBABLY
 CC INVOLVED IN CELL CONTACT.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;
 CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M88710; AAA40859.1; -;
 DR EMBL; M88711; AAA40860.1; -;
 DR EMBL; M88709; AAA40858.1; -;
 DR PIR; JCI239; JCI239;
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00408; Igc2; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
 KW Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 27
 FT CHAIN 1 322
 FT PROPEP 323 345
 FT DOMAIN 50 122
 FT DOMAIN 150 209
 FT DOMAIN 237 303
 FT DISULFD 57 115
 FT DISULFD 157 202
 FT DISULFD 244 296
 FT CARBOHYD 44 44
 FT CARBOHYD 70 70
 FT CARBOHYD 140 140
 FT CARBOHYD 285 285
 FT CARBOHYD 293 293
 FT CARBOHYD 306 306
 FT LIPID 322 322
 FT VARSPLIC 1 27
 SQ SEQUENCE 345 AA; 38067 MW; A3181B0753F9658E CRC64;
 Query Match 6.2%; Score 128; DB 1; Length 345;
 Best local Similarity 25.3%; Pred. No. 0.0041;
 Matches 73; Conservative 35; Mismatches 110; Indels 70; Gaps 17;
 QY 3 RHLLVPEAVSGSGGNEVI-EGPQNAVTLKSGQARFNCTVSGMKLIMWALSDMWVL---58
 DB 20 RLLELVPPGVRSQDAPFPKAMDVTVRQGBSATLRLCTIDRTYTRVM-LNRSTIYLAG 78
 QY 59 ---SVRMEPIITNDRTSQRYDQGNFTSEMIITHNPEPSDGNIRCSLO-----NSR 108
 DB 79 NDKWSIDRVLITLVN---TPQY-----SIMQNDVYDEGYTCSQVTDNHPKSR 127
 QY 109 LHGSALVYQWGEFLFISVNLVVAENEPEYTC---PSHWRLPLDISLGLVSHS 164
 DB 128 VH---LIVQVPPQIMNISDITVNEISVLLCLAIARPE---PVTWR-----HL 172
 QY 165 SYV---FVPEPSDLSQASVSLALPQSGTLTCAVATKSKLARKSATVNLVTR-----IR 215
 DB 173 SVKEGQGFVSEDEYLE---ISDIKRDQSGEYESAL-NDVAADPVKRVKIVNTPIYS 227

QY 216 CPDITGGGINIPGVL---SSLPISLQFSLEPTWCK-----VGLGLAGTML 255
 DB 228 KAKNTGSVQKGIILSCFASAVPMAEFQ---WFKEDTRIATGLDGVRI 272
 RESULT 9
 ID DCC_HUMAN STANDARD; PRT; 1447 AA.
 AC P43146;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor suppressor protein DCC precursor (colorectal cancer suppressor).
 GN DCC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-95011532; PubMed-7926722;
 RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
 RA Vogelstein B.;
 RT "The DCC gene product in cellular differentiation and colorectal
 RT tumorigenesis.";
 RL Genes Dev. 8:1174-1183(1994).
 RN [2]
 RP SEQUENCE OF 1-750 FROM N.A.
 RA MEDLINE-90100559; PubMed-2294591;
 RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
 RA Vogelstein B.;
 RT "Identification of a chromosome 18q gene that is altered in
 RT colorectal cancers.";
 RL Science 247:49-56(1990).
 RN [3]
 RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).
 RA MEDLINE-91121517; PubMed-1991322;
 RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
 RA Oliner J.D., Kinzler K.W., Vogelstein B.;
 RT "Scrambled exons.";
 RL Cell 64:607-613(1991).
 RN [4]
 RP GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.
 RA MEDLINE-94245241; PubMed-8188295;
 RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
 RT "The DCC gene: structural analysis and mutations in colorectal
 RT carcinomas.";
 RL Genomics 19:525-531(1994).
 RN [5]
 RP VARIANT CARCINOMA THR-168, AND VARIANT GHY-201.
 RA MEDLINE-94243823; PubMed-8187090;
 RA Miyake S., Nagai K., Yoshino K., Ota M., Endo M., Yuasa Y.;
 RT "Point mutations and allelic deletion of tumor suppressor gene DCC in
 RT human esophageal squamous cell carcinomas and their relation to
 RT metastasis.";
 RL Cancer Res. 54:3007-3010(1994).
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; X76132; CAAS3735.1; -
 DR EMBL; M32292; AAA5751.1; -
 DR EMBL; M32286; AAA52174.1; -
 DR EMBL; M32288; AAA52175.1; ALT_SEQ.
 DR EMBL; M32280; AAA52176.1; -
 DR EMBL; M63696; AAA52177.1; -
 DR EMBL; M63700; AAA52178.1; -
 DR EMBL; M63702; AAA52179.1; -
 DR EMBL; M63718; AAA52180.1; -
 DR EMBL; M63698; AAA52181.1; -
 DR PIR; A54100; A54100.
 DR PIR; A40098; A40098.
 DR PIR; A38442; A38442.
 DR HSP; P56276; 1TLK.
 DR GeneW; HGNC:2701; DCC.
 DR MIM; 120470; -
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR003962; FN.II_repeat.
 DR InterPro; IPR003006; 19_LMC.
 DR InterPro; IPR003598; 19_C2.
 DR InterPro; IPR003600; 19_Like.
 DR Pfam; PF00041; fn3; 6.
 DR Pfam; PF00047; 1g; 4.
 DR PRINTS; PR00014; FNTYPE.III.
 DR SMART; SM00060; FN3; 6.
 DR SMART; SM00410; IG_Like; 2.
 DR SMART; SM00408; IGC2; 3.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT DOMAIN 26 1097 TUMOR SUPPRESSOR PROTEIN DCC.
 FT TRANSMEM 1098 1122 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1123 1447 POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 154 124 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 154 219 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 354 317 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 345 407 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.
 FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.
 FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.
 FT DISULFID 61 117 BY SIMILARITY.
 FT DISULFID 161 212 BY SIMILARITY.
 FT DISULFID 261 310 BY SIMILARITY.
 FT DISULFID 352 400 BY SIMILARITY.
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 168 168 M -> T (IN OESOPHAGEAL CARCINOMA).
 FT VARIANT 201 201 /FTID-VAR_003909.
 FT VARIANT 201 201 R -> G.
 FT VARIANT 1375 1375 /FTID-VAR_003910.
 FT VARIANT 1375 1375 P -> H (IN A COLORECTAL CARCINOMA).
 FT CONFLICT 138 138 /FTID-VAR_003911.
 FT CONFLICT 233 329 MISSING (IN REF. 3).
 FT CONFLICT 421 421 MISSING (IN REF. 3).
 SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766BD0471F CRC64;
 Query Match 6.2%; Score 128; DB 1; Length 1447;

Best Local Similarity 24.8%; Pred. No. 0 024;
 Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SCGSENEV-----IEGPQATVTKGSGARNCYVSGW--KLIMNLSDM 55
 Db 220 SRTGNEAEVRLISDPGLRHLRYFLQRPNSVVAIEGKDAVLECCVS--GYPPSFYTLRGE 278

QY 56 VVLSVRMEPIITNDRTSORYDOGGNFTSEMIIHNVEPSDGNINC--SLQSRRLGSA 113
 Db 279 VI-----QURSKYSLG--SNLISVTDDSDMYTCVYTKNENISASA 323

QY 114 YLVQVNGELFIPSVNLVAENEPCEVTCPSHWTRLPDISW--ELGLVSHSSYFVPEP 172
 Db 324 ELFTLVPPWFLNHPNSNLVAYESMDIEFCVVS--GKPYPTVNMKNKDVPYIPSPYFQI 382

QY 173 SDLOSANVSIILATPQNSGTLTCVATKSLAKRSATVNLTVICPDGTGGGINIPGVLS 232
 Db 383 SNLR---ILGVKSDGEFYCVAENAGNAGTSADLIYKPAIPSSS-----VLPS 430

QY 233 LP 234
 Db 431 AP 432

RESULT 10
 KMLS_HUMAN STANDARD; PRT: 1914 AA.
 AC 015746; Q14844; Q16794; Q95796; Q95797; Q95798; Q95799; Q9UBG5;
 AC Q9UTY9;
 DT 15-JUN-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin light chain kinase, smooth muscle and non-muscle isozymes
 DE (EC 2.7.1.117) (MCK) [contains: Telokin (kinase related protein)]
 DE (KRP).
 GN MYLK OR MLCK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=umbilical vein endothelial cells;
 RX MEDLINE=97304466; Pubmed=9160829;
 RA Garcia J.G.N., Lazar V.L., Gilbert-McClain L.I., Gallagher P.J.,
 RA Verin A.D.;
 RT "Myosin light chain kinase in endothelium: molecular cloning and
 RT regulation.";
 RL Am. J. Respir. Cell Mol. Biol. 16:489-494(1997).
 RN [2]
 RP REVISIONS.
 RA Birukov K.G., Garcia J.G.N.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3a; 3B AND 4).
 RC TISSUE=umbilical vein;
 RX MEDLINE=99216419; Pubmed=10198165;
 RA Lazar V.L., Garcia J.G.N.;
 RT "A single human myosin light chain kinase gene (MLCK; MYLK).";
 RL Genomics 57:256-267(1999).
 RN [4]
 RP REVISIONS (ISOFORM 2).
 RA Birukov K.G., Garcia J.G.N.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 923-1914 FROM N.A.
 RC TISSUE=hippocampus;
 RX MEDLINE=96121365; Pubmed=8575746;
 RA Potier M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossier J.,
 RA Turnell W.G.;
 RT "The human myosin light chain kinase (MLCK) from hippocampus:
 RT cloning, sequencing, expression, and localization to 3qcen-q21.";
 RL Genomics 29:562-570(1995).

[6]
 RN SEQUENCE OF 1614-1914 FROM N.A.
 RP TISSUE-Lung, and Placenta;
 RX MEDLINE=20007838; PubMed=10536370;
 RA Waterson D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,
 RA Shinsky V.P., Van Eldik L.J., Halech J.;
 RT "Analysis of the kinase-related protein gene found at human chromosome
 RT 3q21 in a multi-gene cluster: organization, expression, alternative
 RT splicing and polymorphic marker";
 RL J. Cell. Biochem. 75:481-491(1999).
 RN [7]
 RP SEQUENCE OF 1456-1914 FROM N.A.
 RC TISSUE-Placenta;
 RA submitted M.D.;
 RL submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR
 CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE
 CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT
 CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL
 CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE
 CC EDEMA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO
 CC CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE
 CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED
 CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT
 CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.
 CC -1- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
 CC light-chain] phosphate.
 CC -1- SUBUNIT: TELOKIN BINDS CALMODULIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS, A NON-MUSCLE FORM (THE
 CC LONGEST FORM), A SMOOTH-MUSCLE FORM AND TELOKIN (A C-TERMINAL
 CC SECTION WITH NO CATALYTIC ACTIVITY); ARE PRODUCED BY ALTERNATIVE
 CC INITIATION.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: 1 (SHOWN HERE), 2, 3A,
 CC 3B, 4 AND DEL-1790; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES ARE
 CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN
 CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE
 CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS
 CC THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN
 CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.
 CC -1- PFM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: UA8959: AAC18423.2; -
 DR EMBL: AF069601: AAD15921.2; -
 DR EMBL: AF069602: AAD15922.1; -
 DR EMBL: AF069603: AAD15923.1; -
 DR EMBL: AF069604: AAD15924.1; -
 DR EMBL: X85337: CAAS9685.1; -
 DR EMBL: AF096771: AAD51380.1; -
 DR EMBL: AF096766: AAD51380.1; JOINED.
 DR EMBL: AF096767: AAD51380.1; JOINED.
 DR EMBL: AF096768: AAD51380.1; JOINED.
 DR EMBL: AF096769: AAD51380.1; JOINED.
 DR EMBL: AF096770: AAD51380.1; JOINED.
 DR EMBL: AF096771: AAD51380.1; JOINED.
 DR EMBL: AF096774: AAD54018.1; -
 DR EMBL: AF096771: AAD51381.1; -
 DR EMBL: AF096769: AAD51381.1; JOINED.
 DR EMBL: AF096770: AAD51381.1; JOINED.
 DR HSSP: O63450: 1A06.
 DR GeneW: HGNC:7590; MYLK.
 DR MIM: 600922; -
 DR InterPro: IPR000719; Euk.pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003600; Ig_1like.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PR00041; fn3; 1.
 DR Pfam: PR00047; Ig; 8.
 DR Pfam: PR00069; pkinase; 1.
 DR PRINTS: PR00014; FNTYPEIII.
 DR ProDom: PD000001; Euk.pkinase; 1.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00410; Ig_1like; 1.
 DR SMART: SM00408; ICG2; 8.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; phosphorylation; immunoglobulin domain; Repeat;
 KW Alternative initiation; 1914
 FT CHAIN 923 1914
 FT
 FT CHAIN 1761 1914
 FT INIT_MET 923 923
 FT
 FT INIT_MET 1761 1761
 FT DOMAIN 1343 1413
 FT DOMAIN 1464 1719
 FT DOMAIN 1711 1774
 FT DOMAIN 1824 1891
 FT NP_BIND 1470 1478
 FT BINDING 1493 1493
 FT ACT_SITE 1585 1585
 FT DOMAIN 1906 1914
 FT DOMAIN 868 998
 FT REPEAT 868 885
 FT REPEAT 896 923
 FT REPEAT 924 951
 FT REPEAT 952 979
 FT REPEAT 980 998
 FT DOMAIN 999 1063
 FT REPEAT 999 1003
 FT REPEAT 1004 1015
 FT REPEAT 1016 1027
 FT REPEAT 1028 1039
 FT REPEAT 1040 1051
 FT REPEAT 1052 1063
 FT VARSPLIC 437 506
 FT
 FT VARSPLIC 1433 1439
 FT
 FT VARSPLIC 1473 1546
 FT
 FT VARSPLIC 1655 1705
 FT VARSPLIC 1790 1790
 FT CONFLICT 933 933
 FT CONFLICT 963 963
 FT CONFLICT 1022 1022
 FT CONFLICT 1048 1050
 FT CONFLICT 1162 1162
 FT CONFLICT 1210 1210
 FT CONFLICT 1280 1280
 FT CONFLICT 1284 1284
 FT
 FT CONFLICT 1300 1300
 FT CONFLICT 1316 1316
 FT CONFLICT 1326 1326
 FT
 A -> G (IN REF. 5).
 L -> S (IN REF. 5).
 T -> S (IN REF. 5).
 VSGIKPEVAMFLEGTPEPRREGSIEYEDGSHYLLKLA
 RTROSGTISGTASNAQGVSCSWTLQVER -> G (IN
 ISOFORM 2 AND ISOFORM 3B).
 DEVEYSD -> MKMKCQT (IN ISOFORM 3A,
 ISOFORM 3B AND ISOFORM 4).
 GKFGQVPELVKTKRQVAKGKFFKAYSASKEENIROELISM
 NCIIHKPRKQCVDAPEERAKIVMVLIVSGSEL -> L
 (IN ISOFORM 4).
 MISSING (IN ISOFORM 3A AND ISOFORM 3B).
 MISSING (IN ISOFORM DEL-1790).
 V -> M (IN REF. 5).
 S -> P (IN REF. 3; AAD15922).
 P -> A (IN REF. 5).
 KPM -> EAH (IN REF. 5).
 P -> L (IN REF. 3; AAD15922/AAD15923).
 L -> P (IN REF. 5).
 E -> D (IN REF. 3; AAD15922/AAD15923).
 M -> I (IN REF. 3; AAD15922/AAD15923/
 AAD15924).

FT CONFLICT 1478 1478 V -> C (IN REF. 5).

FT CONFLICT 1511 1511 S -> T (IN REF. 3); AADI5922/AADI5923).

FT CONFLICT 1563 1563 I -> T (IN REF. 5).

FT CONFLICT 1609 1609 A -> P (IN REF. 5).

FT CONFLICT 1639 1639 G -> R (IN REF. 5).

FT CONFLICT 1639 1639 G -> D (IN REF. 3); AADI5922/AADI5923/

FT CONFLICT 1639 1639 AADI5924).

Query Match 6.2% Score 128; DB 1; Length 1914;

Best Local Similarity 20.3%; Pred. No. 0.034;

Matches 86; Conservative 56; Mismatches 149; Indels 130; Gaps 19;

24 PONATVLKSGQARFNCVSGM--KLIMWALSDWVLSVRPMEITITNDRETSQRYDGG 81

39 PNLCKIKGATAKFEGRY--RGYPQVYTW-----RNGQPTSGRFL--LDCGI 85

82 NPTSMITHNVPDSGNCIRSLNSRLHSAVLTVOY-----MGL 123

86 RGTFSVLVHAHEEDRGKYTCENTNG--SGARQVVELTVESGSPAKOLGQPVVSKTLGDR 143

124 FL-----PSY-----NLVVAENE-----PCEVTCPLPSHWTRLPDISWELG- 158

144 FSASAVETRPSTMGECPPKFKATKGRVYVKEGMRSCKITGRPO-----PQYTNLKN 198

159 LLVSHSYFYVPEPSDLOSAYSIALTPQSNGLTLCVATFMSKLARKSA-----TVN 210

199 VLPQSAVARSYSEKNGM--VLEIHGVNODDVGYTCLVAVNSGKASMSAELSIQGLDSAN 257

211 LTVTCGPDITGGI-----NTPGVLSLPISLGFSLPTMGKVGGLAGIMLLTPTCLTIR 265

258 RSEVETATNSDVAKREVTNYSKESKIDSL----- 288

266 CCCCRRCGCCGCCRCRCRKRFRIOFKSEKKTETETESGENE--SGYNS 322

289 -----BAAKSNKNCSSPQGGSPPMANQPPR-RESLTESCXKSPPTAPQPT 336

323 DEQKTTETASLPKSCSSSDPEQRNSCGPQHQRADQR-----PPRPASPASTNLASPE 378

337 VQKTTSSITLQARVQ--PEPRAPGLGVLPSPGSEKRRAPRPPATFPTROPGLGSQD 393

379 KVS 381

394 VVS 396

DB 394 VVS 396

RESULT 11

PKTK_HUMAN STANDARD: PRT: 1070 AA.

AC Q13508; Q13417;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tyrosine-protein kinase-like 7 precursor (colon carcinoma kinase-4) (CCR-4).

GN PTK7 OR CCK4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Colon carcinoma, and placenta;

RX MEDLINE=96074849; PubMed=7478540;

RA Mosse K., Jallat B., Alves F., Sures I., Plowman G.D., Ullrich A.,

RT "Colon carcinoma kinase-4 defines a new subclass of the receptor

RT tyrosine kinase family."

RL Oncogene 11:2179-2184(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-Fibroblast;

RX MEDLINE=97037064; PubMed=8882711;

RA Park S.-K., Lee H.-S., Lee S.-T.,

RT "Characterization of the human full-length PTK7 cDNA encoding a

RT receptor protein tyrosine kinase-like molecule closely related to

RT chick KLG.*;

RL J. Biochem. 119:235-239(1996).

RN [3]

RP SEQUENCE FROM N.A., AND REVISION TO 834.

RA Lee S.-T., Park S.-K., Lee H.-S., Ji A.R., Jung J.W.,

RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY

CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE

CC PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR

CC PROGRESSION MARKER.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, PANCREAS,

CC KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID

CC GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN

CC ERYTHROLEUKEMIA CELLS. BUT NOT EXPRESSED IN COLON.

CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN

CC RECEPTOR SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: U33635; AAA87565.1; -

DR EMBL: U40271; AAC50484.2; -

DR EMBL: AF447176; AAL39062.1; -

DR EMBL: AF447157; AAL39062.1; JOINED.

DR EMBL: AF447158; AAL39062.1; JOINED.

DR EMBL: AF447162; AAL39062.1; JOINED.

DR EMBL: AF447164; AAL39062.1; JOINED.

DR EMBL: AF447167; AAL39062.1; JOINED.

DR EMBL: AF447170; AAL39062.1; JOINED.

DR EMBL: AF447171; AAL39062.1; JOINED.

DR EMBL: AF447173; AAL39062.1; JOINED.

DR EMBL: AF447174; AAL39062.1; JOINED.

DR EMBL: AF447175; AAL39062.1; JOINED.

DR HSSP: P08631; IAD5.

DR Genew: HGNC:9618; PTK7.

DR MTM: 601890; -

DR InterPro: IPR000719; Euk_pkinase.

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_Like.

DR InterPro: IPR002011; RTKinaseII.

DR InterPro: IPR001245; Tyr_pkinase.

DR Pfam: PF00047; Ig_7.

DR Pfam: PF00069; pkinase_1.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Euk.pkinase; 1.

DR SMART: SM00410; Ig_Like; 2.

DR SMART: SM00408; IgC2; 5.

DR SMART: SM00219; TYRK; 1.

DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; FALSE NEG.

KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;

KW Immunoglobulin domain; Repeat.

FT SIGNAL 1 30

FT CHAIN 31 1070

FT DOMAIN 31 704

FT TRANSMEM 705 725

FT DOMAIN 726 1070

FT DOMAIN 46 108

FT DOMAIN 143 208

FT DOMAIN 239 308

FT DOMAIN 336 398

FT DOMAIN 426 488

FT DOMAIN 517 577

POTENTIAL.

TYROSINE-PROTEIN KINASE-LIKE 7.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

IG-LIKE C2-TYPE DOMAIN 1.

IG-LIKE C2-TYPE DOMAIN 2.

IG-LIKE C2-TYPE DOMAIN 3.

IG-LIKE C2-TYPE DOMAIN 4.

IG-LIKE C2-TYPE DOMAIN 5.

IG-LIKE C2-TYPE DOMAIN 6.

FT DOMAIN 606 671 IG-LIKE C2-TYPE DOMAIN 7.
 FT DOMAIN 796 1066 PROTEIN KINASE; INACTIVE.
 FT DISULFID 53 101 BY SIMILARITY.
 FT DISULFID 150 200 BY SIMILARITY.
 FT DISULFID 246 301 BY SIMILARITY.
 FT DISULFID 343 391 BY SIMILARITY.
 FT DISULFID 433 481 BY SIMILARITY.
 FT DISULFID 524 570 BY SIMILARITY.
 FT CARBOHYD 613 664 BY SIMILARITY.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC...)
 FT CARBOHYD 175 175 N-LINKED (GLCNAC...)
 FT CARBOHYD 184 184 N-LINKED (GLCNAC...)
 FT CARBOHYD 214 214 N-LINKED (GLCNAC...)
 FT CARBOHYD 268 268 N-LINKED (GLCNAC...)
 FT CARBOHYD 283 283 N-LINKED (GLCNAC...)
 FT CARBOHYD 405 405 N-LINKED (GLCNAC...)
 FT CARBOHYD 463 463 N-LINKED (GLCNAC...)
 FT CARBOHYD 567 567 N-LINKED (GLCNAC...)
 FT CARBOHYD 646 646 N-LINKED (GLCNAC...)
 FT CARBOHYD 92 92 P -> R (IN REF. 2 AND 3).
 FT CONFLICT 147 147 K -> T (IN REF. 2 AND 3).
 FT CONFLICT 207 207 S -> G (IN REF. 2 AND 3).
 FT CONFLICT 495 495 VL -> RV (IN REF. 2 AND 3).
 FT CONFLICT 515 515 E -> E (IN REF. 2 AND 3).
 FT CONFLICT 881 881 G -> G (IN REF. 2 AND 3).
 FT CONFLICT 969 969 A -> P (IN REF. 2 AND 3).
 FT CONFLICT 992 992 S -> F (IN REF. 2 AND 3).
 SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B8E3698A5 CRC64;

Query Match 6.0%; Score 125; DB 1; Length 1070;
 Best Local Similarity 24.8%; Pred. No. 0.029;
 Matches 59; Conservative 34; Mismatches 109; Indels 36; Gaps 10;

QY 11 AVGSSGNEVEYEPONATVLSQARFNCVTS-OGMKLIMWALSDWVSVRPMELITN 69
 Db 218 SINDESARVAVLPQDVVARYEAMFHCFSQAPPSPLOMFEDETPIITNSRPHLR 277
 QY 70 DRTSQRIDGNGFTSEMIHNVPSDSGNIRCSLONS- LHSAYLVTVQVGELE 124
 Db 278 ATVFA-----NCSLLTQVRPNRAGIYRCIGOGKRPPIILBATHLAEIEMDPLF 328
 QY 125 IPSVNLVVAENPEVTCIPSHWTRLPDISME-LGL-LVSHSYTVVPPESDQSAVSL 182
 Db 329 EPRVFTAGSE---RYTCPLPKLPSPVSWHEHAGVRLPHGRVY-----QKGHELV 378
 QY 183 ALTPQSN-GLTLCVATWKSLSKARSATVNLV-----IRCPDGTGGGINPCVLSL 233
 Db 379 ANTAESDAGVYTCMA--NLAGGRDQVNTVATVPSWLKKPDQSLDEGKREYLDCL 434

RESULT 12
 PGHM_HUMAN STANDARD; PRT; 4393 AA.
 AC P98160; 016287;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Basement membrane-specific heparan sulfate proteoglycan core
 DE protein precursor (HSPG) (Perlecan) (PLC).
 GN HSPG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92112994; PubMed=1730766;
 RA Kallunki P., Tryggvason K.;
 RT "Human basement membrane heparan sulfate proteoglycan core protein: a
 RT 467-kD protein containing multiple domains resembling elements of the
 RT low density lipoprotein receptor, laminin, neural cell adhesion
 RT molecules, and epidermal growth factor.";
 RL J. Cell Biol. 116:559-571(1992).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin, and Colon;
 RX MEDLINE=92235084; PubMed=1569102;
 RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
 RT "Primary structure of the human heparan sulfate proteoglycan from
 RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
 RT domain homologous to the low density lipoprotein receptor, laminin,
 RT neural cell adhesion molecules, and epidermal growth factor.";
 RL J. Biol. Chem. 267:8544-8557(1992).
 RN [3]
 RP SEQUENCE OF 1018-1472 FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=91365373; PubMed=1679749;
 RA Dodge G.R., Kovacszy I., Chu M.L., Hassell J.R., McBride O.W.,
 RA Yi H.F., Iozzo R.V.;
 RT "Heparan sulfate proteoglycan of human colon: partial molecular
 RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
 RT short arm of human chromosome 1.";
 RL Genomics 10:673-680(1991).
 RN [4]
 RP SEQUENCE OF 892-1398 FROM N.A.
 RC TISSUE=Fibroblasts;
 RX MEDLINE=92120660; PubMed=1685141;
 RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
 RA Tryggvason K.;
 RT "Cloning of human heparan sulfate proteoglycan core protein,
 RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of
 RT a BamHI restriction fragment length polymorphism.";
 RL Genomics 11:389-396(1991).
 RN [5]
 RP SEQUENCE OF 1-21 FROM N.A.
 RX MEDLINE=94052171; PubMed=8234307;
 RA Cohen I.R., Graessle S., Murdoch A.D., Iozzo R.V.;
 RT "Structural characterization of the complete human perlecan gene and
 RT its promoter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
 CC -I- FUNCTION: This protein is an integral component of basement
 CC membranes. It is responsible for the fixed negative electrostatic
 CC charge and is involved in the charge-selective ultrafiltration
 CC properties. It serves as an attachment substrate for cells.
 CC -I- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.
 CC -I- SUBCELLULAR LOCATION: Extracellular.
 CC -I- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
 CC -I- PRT: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
 CC AND O-LINKED OLIGOSACCHARIDES.
 CC -I- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
 CC -I- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.
 CC -I- SIMILARITY: CONTAINS 22 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 SEA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X62515; CAA44373.1; -;
 DR EMBL; M85289; AAA52700.1; -;
 DR EMBL; M64283; AAA52699.1; -;
 DR EMBL; S76436; AAB21121.2; -;
 DR EMBL; L22078; -; NOT_ANNOTATED_CDS.
 DR HSSP; P00740; IEDM.
 DR Siona-ZDPAGE; P98160; -;
 DR Genew; HGNC:5273; HSPG2.

Query	Match	Best local similarity	Score	DB 1	Length
Query 6	LYTPEAVGSG--SGNEVIE-GPQNAITVLKSGOAFNCTVSOG-----KLIIMALSDM 55	62; Conservative	122.5; DB 1; Length 4393;	25.3%; Pred. No. 0.26;	80; Indels 75; Gaps 12
Db 3009	VYVPPSGSSSYRLRSFVSIIDPPSSVYQGDQAFKCIHDGAPISELEKTRNOLEDN 3068				
Query 56	VYLSVPRMEDITINDPFTSQRYDQGNFTSEMIHNHVPSPSGNIR-----SLQ 105				
Db 3069	VHIS--PNSGITI-----LYGRPSNHGTRYKCVASNAIGVADSVY 3106				
Query 106	NSRLHGSAYLTVQVMGELFI--PSVNL-VVAENEPCEVTCLPSHWTRLPDISNEI---- 157				
Db 3107	NLSVHGPRPTVSLPEGPVWVKVGKAVTLECVASGAPRS---SARWTRISSTPAKLHPT 3162				
Query 158	GLLVHSSTYTFYPPESDLSQASVAILTPQSNGLITLVATWTKSLRKRSATVNLVIYIRC 216				
Db 3163	YGLMDHSI-----TVLQISSAKPSDAGTYVCLALNALGTAKQKOVEYIV----- 3204				
Query 217	PQDTG 221				
Db 3205	--DTG 3207				
RESULT 13					
SMP_COTUA	STANDARD; PRT; 620 AA.				
AC 092154;					
DT 15-JUL-1998 (Rel. 36, Created)					
DT 15-JUL-1998 (Rel. 36, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE Schwann cell myelin protein precursor.					
GN SMP.					
OS Coturnix coturnix japonica (Japanese quail).					
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;					
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;					
OC Coturnix					
NCBI_taxid=93934;					

[1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 18-26; 120-132; 135-157 AND
 RP 563-571.
 RX MEDLINE=92153423; PubMed=1739462;
 RA Dulac C., Tropak M.B., Cameron-Curry P., Rossier J., Marshak D.R.,
 Roder J., le Douarin N.M.;
 RT "Molecular characterization of the Schwann cell myelin protein, SMP:
 RT structural similarities within the immunoglobulin superfamily.";
 RL Neuron 8:323-334(1992).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED BY MYELINATING AND
 CC NONMYELINATING SCHWANN CELLS AND OLIGODENDROCYTES.
 CC -1- DEVELOPMENTAL STAGE: FIRST SYNTHESIZED AT EMBRYONIC DAY 5, IT
 CC REMAINS EXPRESSED BY CULTURED SCHWANN CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; S83711; AAB21466.1; -
 DR HSSP; P56276; 1TLK.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00410; Ig_Like; 1.
 DR SMART; SM00408; IgC2; 2.
 KM Myelin: Glycoprotein; Cell adhesion; Transmembrane; Signal; Repeat;
 KW Immunoglobulin domain.
 FT SIGNAL 1 17
 FT CHAIN 18 620
 FT DOMAIN 18 516 SCHWANN CELL MYELIN PROTEIN.
 FT TRANSMEM 517 536 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 537 620 POTENTIAL.
 FT DOMAIN 28 106 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 151 223 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 253 311 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 339 398 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 424 495 IG-LIKE C2-TYPE DOMAIN 3.
 FT DISULFID 35 164 BY SIMILARITY.
 FT DISULFID 40 99 BY SIMILARITY.
 FT DISULFID 158 216 BY SIMILARITY.
 FT DISULFID 260 304 BY SIMILARITY.
 FT DISULFID 346 391 BY SIMILARITY.
 FT DISULFID 420 488 BY SIMILARITY.
 FT DISULFID 431 488 BY SIMILARITY.
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 620 AA; 66943 MW; 004B3EC7EDC18FBA CRC64;

Query Match 5.78; Score 118.5; DB 1; Length 620;
 Best Local Similarity 21.5%; Pred. 0.049;
 Matches 57; Conservative 33; Mismatches 88; Indels 87; Gaps 12;

DB 330 INSLMWVSGDVSVCRAE-----SEPAITLVKGGKVM 365
 QY 183 ALT-----POSNGTLGVATWMSLKRKNTATLVY-----TRQOD 219
 DB 366 AAAYEDHYTMEMRPAREPDGGTYSQVA--ENMGASSTSFNIVPEPLVPSRCTAG 423
 QY 220 TGGGINIFGVLSLP--SIGFSLPT 242
 DB 424 -GDVAVCAVMNSIPDSSLVFELPT 447
 RESULT 14
 ID NEOL CHICK STANDARD; PRT; 1443 AA.
 AC 090610;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neogenin (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=Embryonic brain;
 RX MEDLINE=95105243; PubMed=7806578;
 RA Vielmetter J., Roman J.M., Dreyer W.J.;
 RT "Neogenin, an avian cell surface protein expressed during terminal
 RT neuronal differentiation, is closely related to the human tumor
 RT suppressor molecule deleted in colorectal cancer";
 RL J. Cell Biol. 127:2009-2020(1994).
 CC -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
 CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
 CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
 CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: IN RETINA, EXPRESSED ON GANGLION CELL FIBERS
 CC AS SOON AS THEY BEGIN TO EXTEND THEIR AXONS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
 CC SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U07644; AAC59662.1; -
 DR HSSP; P11276; 2MEN.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam; PF00047; In3; 6.
 DR Pfam; PF00047; In3; 6.
 DR PRINTS; PRO0014; ENTPEPIL1.
 DR SMART; SM00060; FN3; 6.
 DR SMART; SM00410; Ig_Like; 2.
 DR SMART; SM00408; IgC2; 2.
 KM Transmembrane; Immunoglobulin domain; Glycoprotein; Repeat.
 KW NON_TER 1
 FT DOMAIN 1 1090
 FT TRANSMEM 1091 1111 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1112 1443 POTENTIAL.
 FT DOMAIN 1112 1443 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 33 102 IG-LIKE C2-TYPE DOMAIN 1.

```

FT DOMAIN 132 194 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 229 293 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 321 383 IG-LIKE C2-TYPE DOMAIN 4.
FT DOMAIN 422 519 FIBRONECTIN TYPE-III 1.
FT DOMAIN 522 615 FIBRONECTIN TYPE-III 2.
FT DOMAIN 616 714 FIBRONECTIN TYPE-III 3.
FT DOMAIN 720 814 FIBRONECTIN TYPE-III 4.
FT DOMAIN 835 935 FIBRONECTIN TYPE-III 5.
FT DOMAIN 936 1037 FIBRONECTIN TYPE-III 6.
FT DISULFID 139 187 BY SIMILARITY.
FT DISULFID 226 286 BY SIMILARITY.
FT DISULFID 328 376 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 1443 AA; 158050 MW; 5586795579C0E26 CRC64;

Query Match 5.7%; Score 117.5; DB 1; Length 1443;
Best Local Similarity 26.2%; Pred. No. 0.17;
Matches 49; Conservative 28; Mismatches 73; Indels 37; Gaps 11;

OY 21 IECPONATVVKGSQAFNCTVSGW--KLIMVALSDMVLSVRPMPEITNDRTFSORY- 77
DB 218 VQPSSTLTIVTQGNVFP-C-VAGGFPTPVYKTKG-----EELTED---SERFA 264
OY 78 -DQGNFTSEMIITHNVPDSGNIRC-SLONSRLGSAVLTQVNGELFIPSNLVAAE 134
DB 265 LRAGGS----LLISDVEDVGTCTCIADNENETIEAGLELVAVQPEFLKRPANIYAE 320
OY 135 NE----PCEVTCLPSPHWTLPDISW-ELGLVSHSYVPPPSDQSVSLATLPQSN 189
DB 321 SMDVFECEVTGKPT----PTVKWKNQDVVIPSDFYKIVEHNLQ----VLGLVKSDE 371
OY 190 GTLTGVA 196
DB 372 GFYQCIA 378

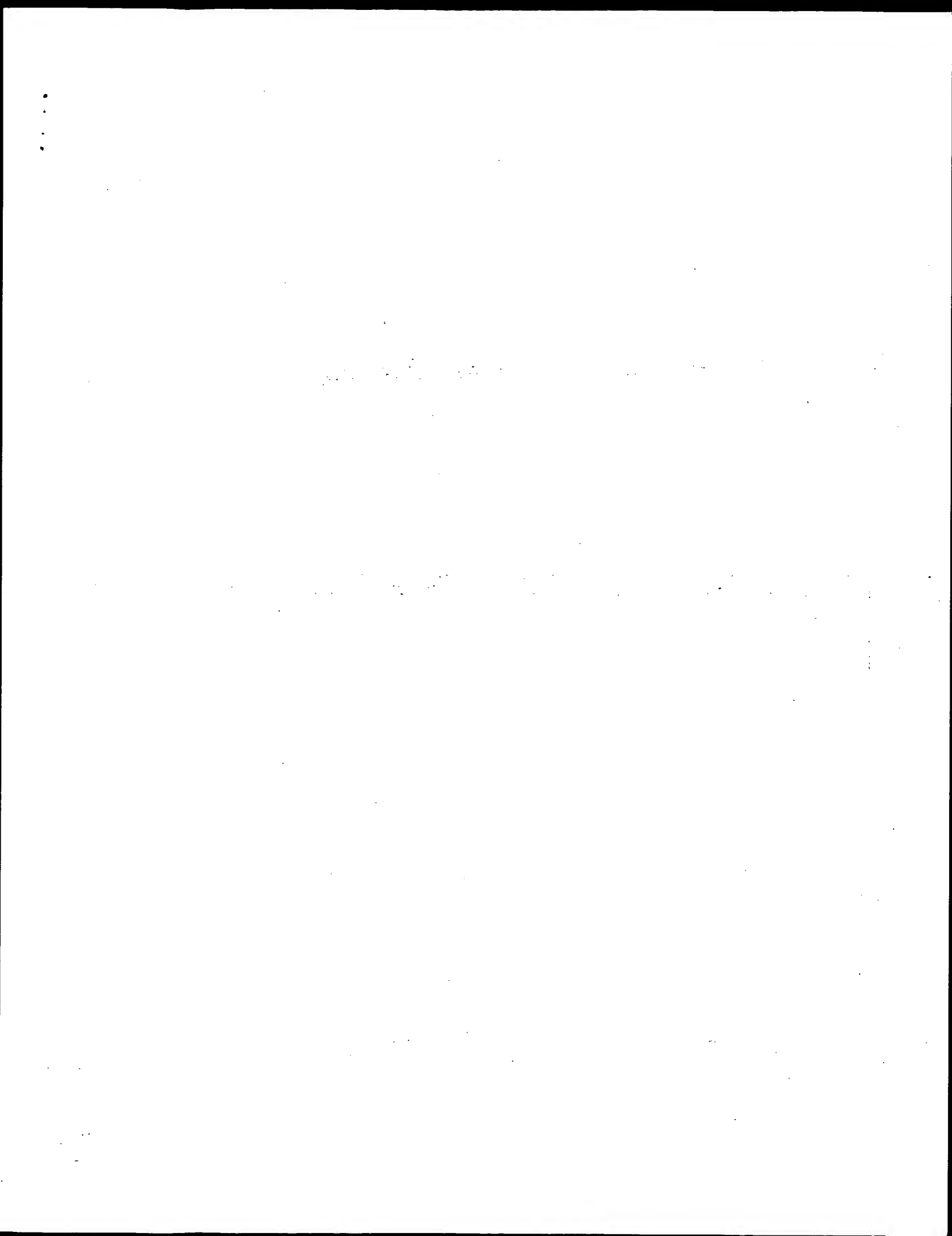
RESULT 15
NEOL_HUMAN
ID NEOL_HUMAN STANDARD; PRT; 1461 AA.
AC 092859; 000340;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neogenin precursor.
GN NEOL OR NGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
OX NCBI_TaxId=9606;
RN [1]
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fetal brain;
RX MEDLINE=97236653; PubMed=9121761;
RT Meyerhardt J.A., Look A.T., Bigner S.H., Fearon E.R.;
RT "Identification and characterization of neogenin, a DCC-related
RT gene.";
RT Oncogene 14:1129-1136(1997).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fetal brain;
RX MEDLINE=97312693; PubMed=9169140;
RX Vielmetter J., Chen X.-N., Miskewich F., Lane R.P., Yamakawa K.,
RX Korenberg J.R., Dreyer W.J.;
RT "Molecular characterization of human neogenin, a DCC-related protein,
RT and the mapping of its gene (NEOL) to chromosomal position 15q22.3-

```

```

RT 923.";
RL Genomics 41:414-421(1997).
CC -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS, 1 (SHOWN HERE) AND 2;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED AND ALSO IN CANCER CELL
CC LINES.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; U61262; AB017263.1; -
DR EMBL; U72391; AAC51287.1; -
DR HSPB; P02751; ITTF.
DR GeneW; HGNC:754; NEOL.
DR MIM; 601907; -
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII repeat.
DR InterPro; IPR003062; Ig_MHC.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_1like.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; Ig; 4.
DR PRINTS; PR00014; ENTPEIII.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00410; IG_1like; 1.
DR SMART; SM00408; IGC2; 3.
DR Transmembrane; Immunoglobulin domain; Glycoprotein; Signal; Repeat;
KW Alternative splicing.
FT SIGNAL 1 33
FT CHAIN 34 1461
FT DOMAIN 34 1105
FT TRANSMEM 1106 1126
FT DOMAIN 1127 1461
FT DOMAIN 67 136
FT DOMAIN 166 228
FT DOMAIN 263 327
FT DOMAIN 355 417
FT DOMAIN 436 533
FT DOMAIN 536 629
FT DOMAIN 630 729
FT DOMAIN 735 829
FT DOMAIN 850 950
FT DOMAIN 951 1052
FT DOMAIN 1118 1121
FT DISULFID 74 129
FT DISULFID 173 221
FT DISULFID 270 320
FT DISULFID 362 410
FT CARBOHYD 73 73
FT CARBOHYD 210 210
FT CARBOHYD 326 326
FT CARBOHYD 470 470
FT CARBOHYD 489 489
FT CARBOHYD 639 639
FT CARBOHYD 715 715
FT CARBOHYD 909 909
FT CARBOHYD 1248 1300
FT VARSPLIC 168 168
FT CONFLICT 1461 AA; 159958 MW; 7AAE897E6965A21 CRC64;
SO SEQUENCE

```

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 18:09:54 ; Search time 19.0659 Seconds
(without alignments)
1946.300 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077

Sequence: 1 MERHLLTPEAVGSSGSGNEV.....HPQASFNLASPEKVSNTTV 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	7.3	404	1	I61596 advanced glycosyla
2	139.5	6.7	344	2	I56551 neurotlimin - rat
3	133	6.4	1427	2	I51659 tumor suppressor
4	130	6.3	345	2	JC4025 oploid-binding cel
5	129.5	6.2	164	2	T24272 hypothetical prote
6	129	6.2	345	2	S03197 oploid-binding pro
7	129	6.2	6642	2	T29757 protein UNC-89 - C
8	128	6.2	345	2	JC1239 oploid-binding pro
9	128	6.2	1447	2	AS4100 tumor suppressor p
10	127.5	6.1	871	1	I48696 protein-tyrosine k
11	127.5	6.1	881	1	I48697 protein-tyrosine k
12	125	6.0	1070	2	JC4593 protein-tyrosine k
13	122.5	5.9	188	2	T15651 hypothetical prote
14	122.5	5.9	4391	2	A38096 perlecan precursor
15	118.5	5.7	338	2	JC1238 oploid-binding pro
16	118.5	5.7	620	2	JH0593 schwann cell myeli
17	117.5	5.7	1443	2	I50600 neogenin - chicken
18	114.5	5.5	152	2	T18975 hypothetical prote
19	113.5	5.5	416	1	A42879 advanced glycosyla
20	111	5.3	693	2	S49228 sodium-dependent p
21	110.5	5.3	764	2	A49448 irregular chiasm C
22	110.5	5.3	1177	2	T16594 hypothetical prote
23	109.5	5.3	880	1	A53743 protein-tyrosine k
24	109	5.2	391	2	T09058 butyrophilin homol
25	108.5	5.2	487	2	S65133 butyrophilin - mou
26	106.5	5.1	802	2	T13149 mitogen-and stress
27	106.5	5.1	946	1	A47299 ror-related recept
28	106.5	5.1	2491	1	A28372 insulin-like growt
29	106	5.1	1092	1	JN0635 neural cell adhesi

30	105.5	5.1	862	2	I49583 differentiation an
31	105.5	5.1	882	2	I38912 receptor tyrosine
32	105.5	5.1	3707	2	S18252 heparan sulfate pr
33	105	5.1	662	2	T16525 hypothetical prote
34	105	5.1	5825	2	T12117 polyprotein - fava
35	104	5.0	2222	2	T13924 sak protein - fru
36	103.5	5.0	423	2	T29549 hypothetical prote
37	103.5	5.0	1323	2	PN0568 connectin 3b - chi
38	103	5.0	365	2	JC5780 coxsackie- and ade
39	103	5.0	1273	2	T42405 sax-3 protein - Ca
40	102	4.9	721	2	T41530 hypothetical prote
41	102	4.9	868	2	A46512 CD22 homolog/B lym
42	101.5	4.9	344	1	RWRTC2 T-cell surface gly
43	101.5	4.9	421	2	T46266 hypothetical prote
44	101.5	4.9	841	2	JC5894 killer cell inhibi
45	101	4.9	338	2	JC5519 50K glycoprotein p

ALIGNMENTS

RESULT 1
I61596
advanced glycosylation end-products receptor precursor - human
N:Alternate names: advanced glycosylation end product-binding protein, 35K, glycoprot
C:Species: Homo sapiens (man)
C>Date: 24-May-1996 #sequence revision 07-Feb-1997 #text change 16-Jul-1999
C:Accession: I61596; B42879; S27968
R:Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko,
Genomics 23, 408-419, 1994
A:Title: Three genes in the human MHC class III region near the junction with the cia
nterpart of mouse mammary tumor gene Int-3.
A:Reference number: A55562; MUID:95137587; PMID:7835890
A:Accession: I61596
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-404 <RES>
A:Cross-references: GB:D28769; NID:9561657; PIDN:BA05958.1; PID:9561659
R:Neepser, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.;
J Biol Chem 267, 14998-15004, 1992
A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation
A:Reference number: A42879; MUID:92340547; PMID:1378843
A:Accession: B42879
A:Molecule type: mRNA
A:Residues: 1-404 <RES>
A:Cross-references: EMBL:M91211; NID:9190845; PIDN:AA03574.1; PID:9190846
A:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBI:109438)
C:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly
cellular function, thus contributing to tissue lesions in diabetes.
C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C:Genetics:
A:Gene: GDB:AGER
A:Cross-references: GDB:306354; OMIM:600214
A:Map position: 6p21.3-6p21.3
A:Introns: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2
C:Function:
A:Description: neuronal receptor for amphoterin, a DNA-binding protein involved in ne
C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology
C:Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-404/Product: advanced glycosylation end products receptor #status predicted <MAT
F:23-344/Domain: extracellular #status predicted <EXT>
F:31-101/Domain: immunoglobulin homology <IM1>
F:137-210/Domain: immunoglobulin homology <IM2>
F:252-303/Domain: immunoglobulin homology <IM3>
F:345-362/Domain: transmembrane #status predicted <TM>
F:363-404/Domain: intracellular #status predicted <INT>
F:25, 81/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:38-99, 144-208, 255-301/Disulfide bonds: #status predicted

Query Match 7.3%, Score 151, DB 1, Length 404;

Best Local Similarity 23.3%; Pred. No. 0.001;
Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19;

QY 9 PEAVGSGSNEVEIEGPQNAIVLKSGQARFNCVSGW---KLIIMALSDMVLVSRPMEP 65
Db 124 PEIVDSAS--ELTAGVPN-----KVTGCVSGSGYPAGTLISWHDG-----KP 163
QY 66 ITTNDREFTS-----ORYDGGNFT--SEMIITHNVEPSDGNIR---CSLONSRLHGSAY 114
Db 164 LVPNEGVSVKRQTRRHPTGTGLQSLM---VTPARGDPRPFSCSFGSLPRHRL 220
QY 115 LTVQVMGELFIP---SVNLVAENEP-----CEVTCPLSHMTLPDISWE 156
Db 221 RTAPLOPRWVEPPLVEEYOLVY---EPREGAVAPGVTLLICEVPAOPS-----POLHMM 272
QY 157 LGLVSHSSTVPPSPDQSAVSTIALTPQSGNLTLCVATWKSILAKKASATVNTVTRC 216
Db 273 KD-----GVPLPLPPSVLLIPEIGPODQGYSCVAHSHSGPESRAVSISILE- 322
QY 217 PODTGGINIPGVLSLPSLGFSLPTWGVGLAGTMLLT-----PTGTLTRCCCC 269
Db 323 PEEEG-----PTAGSVGSGSLGTLALAGILGIGTALLIGVILM 363
QY 270 RRRCCGCCNCCGCCRCRKRGRFIOFOKSEKERT--NKLETETSGNDSGNSDECKT 327
Db 364 QRR-----ORGERAKAPENOEBEERAEIN----- 389
QY 328 TETASLPKSCSSDPEQNSSCGPP 353
Db 390 -----QSEPEAGESSSTGCP 404

RESULT 2

156551
neurotrophin - rat

C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 19-May-2000
C:Accession: J156551
R:Struyk, A.F.; Canoll, P.D.; Mollfang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.
J. Neurosci. 15, 2141-2156, 1995
A:Title: Cloning of neurotrophin defines a new subfamily of differentially expressed neur
A:Reference number: J156551; MUID:95198094; PMID:7891157
A:Accession: J156551
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <RES>
A:Cross-references: EMBL:U16845; NID:9755184; PIDN:AAA67445.1; PID:9755185
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match

Best Local Similarity 24.9%; Score 139.5; DB 2; Length 344;
Matches 66; Conservative 40; Mismatches 110; Indels 49; Gaps 13;

QY 3 RHLLTPEAVGSGSNEVI-EGPQNAIVLKSGQARFNCVSGWKLIMALSDMVLV-- 59
Db 20 RLFLVPTGVVPSGDAIFPKAMDVTVRQGESATLRCTIDRVTAVMLNKRSTILVAGN 79
QY 60 -----VRPMEPIITNDRFTSQRYDGGNFTSEMIITHNVEPSDGNIRCSLO-----NSRL 109
Db 80 DKWCIDPRVYLLSN---TGTQY-----STEIONVDVDEGPTGCVQTDNHPKTSRV 128
QY 110 HGSAYLTVQVMGELFIPSVNLVAENEPCEVTCPLSHMTLPDISWEGLLVSHSSTYFV 169
Db 129 H-----LTVQVSPKIVEISDISINSGNNISLCTAIGRPE-PLYTWK---HISPRVAHFV 180
QY 170 PEPDQSAVSTIALTPQSGNLTLCVATWKSILAKKASATVNTVTRC-----QDTGGGI 224
Db 181 SEDEYLE-----IOGITRESGGEYCSAS--NDVAAPVRRVAVTVVPIYISEAKGTGVPV 235
QY 225 NIPGVV-----SSLPSLGFSLPTWGR 245
Db 236 GQKGTLOCEASAVPSAEFO---WFK 257

RESULT 3
151669
tumor suppressor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: J151669
R:Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.
Dev. Biol. 166, 654-665, 1994
A:Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in
A:Reference number: J151668; MUID:95113183; PMID:7813784
A:Accession: J151669
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1427 <PIE>
A:Cross-references: EMBL:U10986; NID:9606873; PIDN:AAA70168.1; PID:9606874
C:Genetics:
A:Gene: XDCCA

Query Match

Best Local Similarity 25.7%; Score 133; DB 2; Length 1427;
Matches 59; Conservative 33; Mismatches 90; Indels 48; Gaps 13;

QY 21 IEGPQNAIVLKSGQARFNCVSGW--KLIIMALSDMVLVSRPMEPIITNDRFTSQRYD 78
Db 243 LQPSNVVAIEGDAVLECAVS--GYPTPTVMMQGD-----BPVPIRTR---KYS 288
QY 79 QGGNFTSEMIITHNVEPSDGNIR---SLONSRLHGSAYLTVQVMGELFIPSVNLVAENEP 136
Db 289 VLGG--SNLISNVTDDDAATVCAVTKKENTSPSADLTVMVPPQPLNHPANLVAESM 346
QY 137 PCEVTCI---PSHMTLPDISW-ELGLVSHSSTYFVPEPDSQSAVSTIALTPQSGNLT 191
Db 347 DIEFECAVSGKPS-----PLVTKWTKNGEVVIPSDYQIVDGSMLR---ITGLVKSDEGY 397
QY 192 LTCVA-----TWKSLAKKASATVNTVTRC-RCPQDTGGGINIPGVLS 222
Db 398 YQCIANEAGNIQYQTLIIPDPVAPSSSILPSAPRV-----VPVLVSS 442

RESULT 4

JC4025
opioid-binding cell adhesion protein - human

C:Species: Homo sapiens (man)
C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Mar-2000
C:Accession: JC4025
R:Shark, K.B.; Lee, N.M.
Gene 155, 213-317, 1995
A:Title: Cloning, sequencing and localization to chromosome 11 of a cDNA encoding a h
A:Reference number: JC4025; MUID:95237612; PMID:7721093
A:Accession: JC4025
A:Molecule type: mRNA
A:Residues: 1-345 <SHA>

A:Cross-references: GB:J34774; NID:9514373; PIDN:AAA36387.1; PID:9514374
A:Experimental source: brain
C:Comment: This protein binds opioid alkaloids in the presence of acidic lipids, exhi
C:Genetics:
A:Gene: GDB:OPCML; OBCAM; OPCM
A:Cross-references: GDB:251677; OMIM:600632
A:Map position: 11pter-11qter
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-ter

Query Match

Best Local Similarity 25.6%; Score 130; DB 2; Length 345;
Matches 73; Conservative 34; Mismatches 108; Indels 70; Gaps 17;

QY 3 RHLLTPEAVGSGSNEVI-EGPQNAIVLKSGQARFNCVSGWKLIMALSDMVLV-- 58
Db 20 RLFLVPTGVVPSGDAIFPKAMDVTVRQGESATLRCTIDRVTAVMLNKRSTILVAGN 78
QY 59 -----SVRPMETITNDRFTSQRYDGGNFTSEMIITHNVEPSDGNIRCSLO-----NSR 108
Db 79 NDKWSIDPRVILLVN---TPTQY-----SIMIONVDVDEGPTGCVQTDNHPKTSR 127

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-May-2000
 C/Accession: JCI233
 R:Lippman, D.A.; Lee, N.M.; Loh, H.H.
 Gene 117, 249-254, 1992
 A>Title: Opioid-binding cell adhesion molecule (OBAM)-related clones from a rat brain
 A/Reference number: JCI238; MUID:92347701; PMID:1339369
 A/Accession: JCI233
 A/Molecule type: mRNA
 A/Residues: 1-345 <LIP>
 A/Cross-references: GB:M88710; NID:9203247; PIDN:AAA0859.1; PID:9203248; GB:M88711; NID:9203249
 C/Supplementary: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminal

Query Match 6.2%; Score 128; DB 2; Length 345;
 Best Local Similarity 25.3%; Pred. No. 0.044;
 Matches 73; Conservative 35; Mismatches 110; Indels 70; Gaps 17;

QY 3 RLHLTPAVGSSGNEVI-EGFQNAVTLKSGARFNCVSGWKLIMMALSDMYL--- 58
 Db 20 RLFLVPLTGVPRSGDATPFKAMDNTVROGESATLRCTIDRVTTRVAM-LNRSTILYAG 78
 QY 59 ----SVRPMPEITINDFTSQARDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSR 108
 Db 79 NDKMSIDPRVILVN---IPTQY-----SIMIQNDVYDEGFCYTCVQTDNHPKTSR 127
 QY 109 LHGSAYLVTVGMELFIPSVNLVAENEPCEVTCL-----PSHWTLPLDISMELGLVSHS 164
 Db 128 VH-----LIVGVPPQIMNISSDITVNEISSVTLICLAIQRP-----PLVTWR-----HL 172
 QY 165 SYT-----FVPEPSDLOSANVILALPPQSNGLTCVATWKSLSKARSAVTNLT-----IR 215
 Db 173 SYVEGGGFVSEDEYLE---ISDIKRDQSGEYECAL-NDVAADPVKAVITVWYPIYS 227
 QY 216 CPQDTGGGGINIPGVL-----SLPSLGFSLPTMGK-----VGLGLACTYL 255
 Db 228 KAKNTGVSVGGKILSCASAVPAEFO---WKEDETRLATGLDGVRI 272

RESULT 9

A54100
 Tumor suppressor protein DCC precursor - human
 N/Alternate names: colorectal cancer suppressor DCC
 C/Species: Homo sapiens (man)
 C/Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 05-Nov-1999
 C/Accession: A54100; A40098
 R:Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
 Genes Dev. 8, 1174-1183, 1994
 A>Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.
 A/Reference number: A54100; MUID:95011532; PMID:7926722
 A/Accession: A54100
 A/Molecule type: mRNA
 A/Residues: 1-1447 <HED>
 A/Cross-references: EMBL:X76132; NID:9453209; PIDN:CAA53735.1; PID:9453210
 R:Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamllit
 Science 247, 49-56, 1990
 A>Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.
 A/Reference number: A40098; MUID:90100559; PMID:2294591
 A/Accession: A40098
 A/Molecule type: mRNA
 A/Residues: 1-750 <FEA>
 A/Cross-references: GB:M32292; NID:9181492; PIDN:AAA35751.1; PID:9181493
 C/Genetics:
 A:Gene: GDB:DCC
 A/Cross-references: GDB:119838; OMIM:120470
 A/Map position: 18q21.1-18q21.1
 C/Keywords: transmembrane protein; tumor suppressor
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 6.2%; Score 128; DB 2; Length 1447;
 Best Local Similarity 24.8%; Pred. No. 0.21;
 Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SGSGNEV-----IEGPNAVTLKSGARFNCVSGW- KLIMMALSDM 55
 Db 220 SRGTNEAEVRIILSDPCLHQLTLPSPSVNVAIECKDVLVECCVS-GYPPEFSTWLGE 278
 QY 56 VILSVRPMPEITINDFTSQARDQGNFTSEMIHNVPSDSGNIRCSLQ-----NSR 108
 Db 279 VI-----QLRSKRYSLIG--SNLLISNTYDDDSGMYTCVITYKNENISASA 323
 QY 114 YLTVQVGMELFIPSVNLVAENEPCEVTCCLPSHWTLPLDISW-ELGLVSHSSYFVPEP 172
 Db 324 ELTVLVPPEFLNHPNSLVAIESMDIEFECTVS-GKPPVPMNMKNGDVVIPSDFQIVG 382
 QY 173 SDIQSAVSTLALTPQSNGLTCVATWKSLSKARSAVTNLTVCPODTGGGINIPGVLS 232
 Db 383 SNLR-----ILGVKSKDEGFYCCVAENEGNAGNQSIAQLIVKRAIPSSS-----VLPS 430
 QY 233 LP 234
 Db 431 AP 432

RESULT 10

148696
 Protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 2 - mouse
 N/Alternate names: receptor-type tyrosine kinase
 N/Contents: protein-tyrosine kinase nsk2 precursor, splice form 4
 C/Species: Mus musculus (house mouse)
 C/Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
 R:Ganju, P.; Walls, E.; Brennan, J.; Reith, A.D.
 Oncogene 11, 281-290, 1995
 A>Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase
 A/Reference number: 148696; MUID:9534951; PMID:7624144
 A/Accession: 148696
 A/Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-871 <GANI>
 A/Cross-references: EMBL:X86444; NID:9292723; PIDN:CAA60165.1; PID:9292724
 A/Experimental source: splice form 2
 A/Accession: S60738
 A/Molecule type: DNA
 A/Residues: 1-456, A, 466-871 <GAN2>
 A/Cross-references: EMBL:X86444; NID:9292723
 A/Experimental source: splice form 4
 C/Comment: For alternate splice forms see PIR:148697.
 C/Genetics:
 A:Gene: nsk2
 A/Cross-references: MGI:103308
 C/Supplementary: mouse for-related receptor; immunoglobulin homology; protein kinase ho
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-871/Product: protein-tyrosine kinase nsk2, splice form 2 #status predicted <MAT2
 F:22-456, A, 466-871/Product: protein-tyrosine kinase nsk2, splice form 4 #status pre
 F:135-199/Domain: immunoglobulin homology <IMM1>
 F:226-284/Domain: immunoglobulin homology <IMM2>
 F:498-518/Domain: transmembrane #status predicted <TRM>
 F:575-865/Domain: transmembrane #status predicted <TRM>
 F:583-591/Region: protein kinase ATP-binding motif
 F:222,462/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 6.1%; Score 127.5; DB 1; Length 871;
 Best Local Similarity 21.7%; Pred. No. 0.13;
 Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

QY 19 EVIEGPNAVTLKSGARFNCVSGW- KLIMMALSDMYLSDVSRPMPEITINDFTSQRY 77
 Db 122 KITRPPINVKIIRGLAKAVLPCTTMCNPKRSVYKIGD-----NALRENSRIALE- 171
 QY 78 DQGNFTSEMIHNVPSDSGNIRCSLQNSRLHGSAY-----LTVQVGMELFIPSVNLV 132
 Db 172 -----SGSLRIHNVQEKEDAGYRCVAKNSL--GVAASKVLKLEVEVLGLILAPESHNV 223

Qy 133 AENEPEVTC-----LPSHMTLPDISW-ELGLVSHSSYFVPEPSDLSQASVSLALTPQ 187
 Db 224 TFGSVTLKRCHEIGIP-----VPTISWLENGNNAVSSSGSIQSVNDRVIDSLQLEFTRP 277
 Qy 188 SNGILTCVAT-----WKSLSKARKSATVNLTVIR-----CPQDTGGGINIPG-----V 229
 Db 278 --GLTYCIATNNKGEKFTAKAAATVSIAMWSKQSDQCYCAQYRGEVLMQGPGEKML 335
 Qy 230 LSLPSLIGFSLP-----TWKRVGLGLAGTMTLPTCTLTIRCCCCRRCCGCN-- 277
 Db 336 LVFLPTTSHRDPDAQELLITHTAMNEL-----KAVSPLCRPAABALLCYHLFLCSPG 388
 Qy 278 -----CCRCRC-----FCRC 287
 Db 389 VVPTPMPICREYCLAVKEIFCAK 411

RESULT 11

148697
 protein-tyrosine kinase (EC 2.7.1.112) nsk2 precursor, splice form 1 - mouse
 N:Alternate names: receptor-tyrosine kinase
 M:Contains: protein-tyrosine kinase nsk2 precursor, splice form 3
 C:Species: Mus musculus (house mouse)
 C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
 C:Accession: J4593
 R:Ganju, P.; Wallis, E.; Brennan, J.; Reith, A.D.
 Oncogene 11, 281-290, 1995
 A:Title: Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase
 A:Reference number: 148696, MUID:95349551; PMID:7624114
 A:Accession: 148697
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-881 <GAN1>
 A:Cross-references: EMBL:X86445; NID:9929725; PIDN:CA60166.1; PID:9929726
 A:Experimental source: splice form 1
 A:Accession: 560740
 A:Molecule type: DNA
 A:Residues: 1-456, 'A', 466-881 <GAN2>
 A:Cross-references: EMBL:X86445; NID:9929725
 A:Experimental source: splice form 3
 C:Comment: For alternate splice forms see PIR:148696.
 C:Genetics:
 A:Gene: nsk2
 A:Cross-references: MGI:103308
 C:Superfamily: mouse ror-related receptor; immunoglobulin homology; protein kinase homol
 C:Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein; tyro
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-881/Product: protein-tyrosine kinase nsk2, splice form 1 #status predicted <MAT1>
 F:22-456, 'A', 466-881/Product: protein-tyrosine kinase nsk2, splice form 3 #status predi
 F:42-101/Domain: immunoglobulin homology <IMM1>
 F:135-192/Domain: immunoglobulin homology <IMM2>
 F:226-284/Domain: immunoglobulin homology <IMM3>
 F:498-518/Domain: transmembrane #status predicted <TRM>
 F:575-865/Domain: protein kinase homology <KIN>
 F:583-591/Region: protein kinase ATP-binding motif
 F:222,462/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.1%; Score 127.5; DB 1; Length 881;
 Best Local Similarity 21.7%; Pred. No. 0.13;
 Matches 70; Conservative 41; Mismatches 125; Indels 87; Gaps 16;

Qy 19 EYISGPONATVYKGSQARNCTVSQGW-LIMWALSDWVLSVRPMEITITNDRFTSORY 77
 Db 122 KITRPPIWAKIEGLKAVLPCTWGNPKPSVSWIKGD-----NALRENSRIAALE- 171
 Qy 78 DQGNFTSEMIITHNEPDSGNINCSLQNSRLHGSAY-----LVQWAGLEFISVNLV 132
 Db 172 -----SSLSLHYNQKEDAQYRCVANKNSL--STAYSKLWLEVEVIGRLRAPESHNV 223
 Qy 133 AENEPEVTC-----LPSHMTLPDISW-ELGLVSHSSYFVPEPSDLSQASVSLALTPQ 187
 Db 224 TFGSVTLKRCHEIGIP-----VPTISWLENGNNAVSSSGSIQSVNDRVIDSLQLEFTRP 277

Qy 188 SNGILTCVAT-----WKSLSKARKSATVNLTVIR-----CPQDTGGGINIPG-----V 229
 Db 278 --GLTYCIATNNKGEKFTAKAAATVSIAMWSKQSDQCYCAQYRGEVLMQGPGEKML 335
 Qy 230 LSLPSLIGFSLP-----TWKRVGLGLAGTMTLPTCTLTIRCCCCRRCCGCN-- 277
 Db 336 LVFLPTTSHRDPDAQELLITHTAMNEL-----KAVSPLCRPAABALLCYHLFLCSPG 388
 Qy 278 -----CCRCRC-----FCRC 287
 Db 389 VVPTPMPICREYCLAVKEIFCAK 411

RESULT 12

JC4593
 protein-tyrosine kinase-related receptor PTK7 precursor - human
 N:Alternate names: receptor protein tyrosine kinase-like protein (RPTK)
 C:Species: Homo sapiens (man)
 C:Date: 16-Apr-1996 #sequence_revision 24-May-1996 #text_change 24-Sep-1999
 C:Accession: JC4593
 R:Park, S.K.; Lee, H.S.; Lee, S.T.
 J. Biochem. 119, 235-239, 1996
 A:Title: Characterization of the human full-length PTK7 cDNA encoding a receptor prot
 A:Reference number: JC4593; MUID:97037064; PMID:8882711
 A:Accession: JC4593
 A:Molecule type: mRNA
 A:Residues: 1-1070 <PAR>
 A:Cross-references: GB:U040271; NID:91322231; PIDN:AC50484.1; PID:91322232
 C:Comment: This protein is a member of receptor protein tyrosine kinase family, but p
 C:Genetics:
 A:Gene: GDB:PTK7
 A:Cross-references: GDB:134760; OMIM:601890
 A:Map position: 6p21.1-6p12.2
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: cell adhesion; extracellular protein; glycoprotein; phosphotransferase; t
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-1070/Product: protein-tyrosine kinase 7 #status predicted <MAT>
 F:31-703/Domain: extracellular #status predicted <EXC>
 F:704-725/Domain: transmembrane #status predicted <TRM>
 F:726-1070/Domain: intracellular #status predicted <KIN>
 F:794-1065/Domain: protein kinase homology <KIN>
 F:116,175,184,214,268,283,405,463,567,646/Binding site: carbohydrate (Asn) (covalent)

Query Match 6.0%; Score 125; DB 2; Length 1070;
 Best Local Similarity 24.8%; Pred. No. 0.25;
 Matches 59; Conservative 34; Mismatches 109; Indels 36; Gaps 10;

Qy 11 AVSGSGNEVIEGPONATVYKGSQARNCTVS-QGKMLIMWALSDWVLSVRPMEITITN 69
 Db 218 SIADSEFARVVLAPDQVVARVEAMFHCOFSQAPPSLQWLFEDETPTITNSRPHLRR 277
 Qy 70 DRTSQRYDQGNFTSEMIITHNEPDSGNIRCSLQNSR-----LHGSAYLVQWAGLEF 124
 Db 278 ATYFA-----NSLLTQVPRNAGYRIGGQGRPPIITATLHLAIEDMPLE 328
 Qy 125 IPSVNLVAENEBCEVTCLEPSHMTLPDISWE-LGL-LVSHSSYFVPEPSDLSQASVSL 182
 Db 329 EPRVFTAGSE--RVTCLEPRLKLPESVWMEHAGVRLPHGVY-----QKHELV 378
 Qy 183 ALTPQSN-GILTVATWKSLSKARKSATVNLTV-----TRCPQDTGGGINITGVSSL 233
 Db 379 ANIAESDAGYVTCMA--MLAGORQDVNITVATVPSWMLKRPQDSOLEBGRKPGYLDCL 434

RESULT 13

T15651
 hypothetical protein C27A2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15651
 R:Nhan, M.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid C27A2.

Gene 117, 249-254, 1992

A:Title: Opioid-binding cell adhesion molecule (OBCAM)-related clones from a rat brain c

A:Reference number: JCI238; MUID:92347701; PMID:1339369

A:Accession: JCI238

A:Molecule type: mRNA

A:Residues: 1-338 <LTP>

A:Cross-references: GB:M88709; NID:g203245; PIDN:AAA40858.1; PID:g203246

A:Experimental source: Brain

C:Genetics:

A:Gene: OBCAM

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 5.7%; Score 118.5; DB 2; Length 338;

Best Local Similarity 25.2%; Pred. No. 0.22;

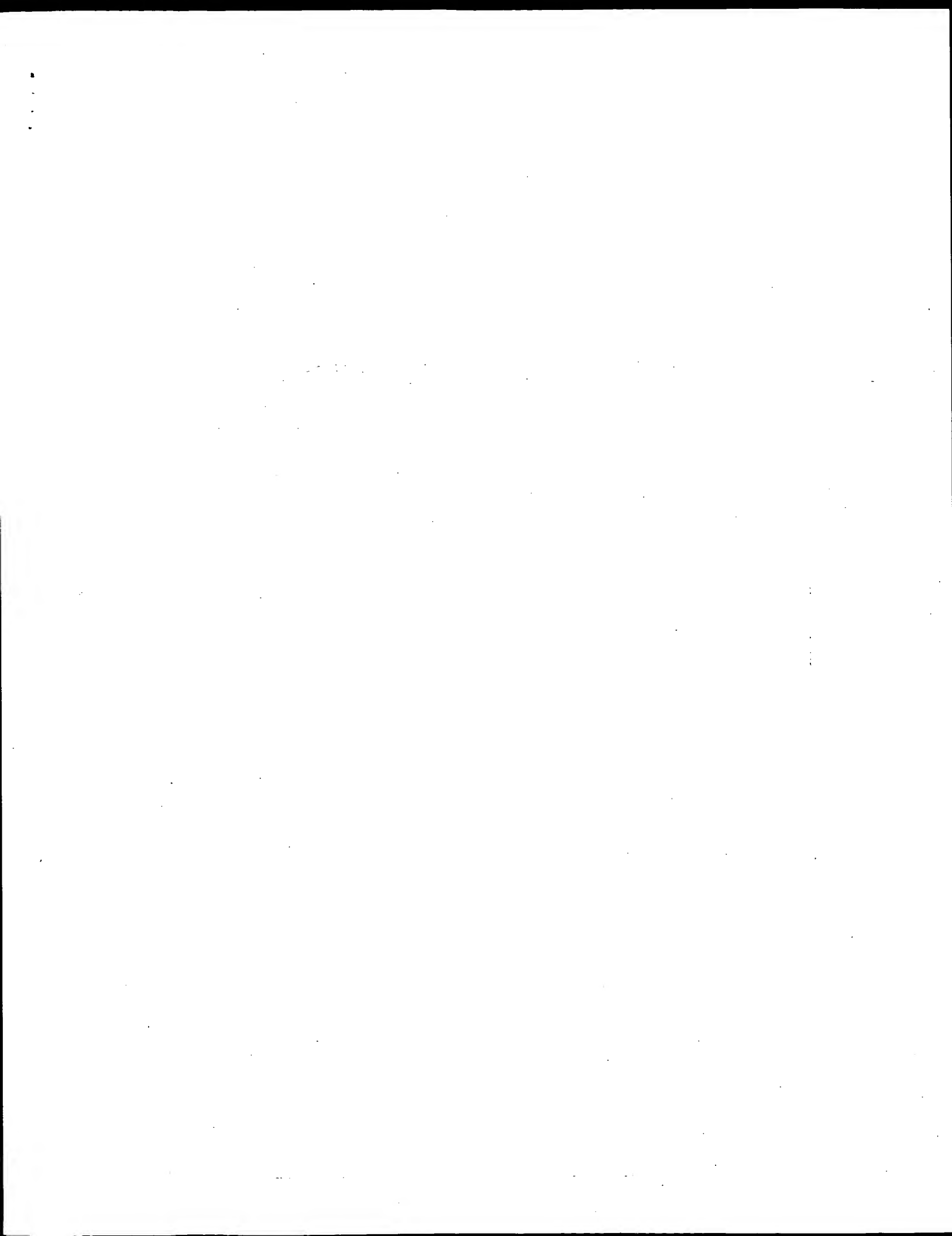
Matches 72; Conservative 36; Mismatches 107; Indels 71; Gaps 18;

```

OY 5  LITVEAVGSSGSGNEVI-EGPQNAFTVLKSGARFNCITYSGWKLIIMALSDMYL----- 58
DB 16 LLETF-GVPRVSGDNTFPKAMDNTVVRQGESATLRCTIIDDRTVRVAM-LNRSTILVAGND 73
OY 59 --SVRPMETITNDRTSQRYDQGNFTSEMIINVEPSDSGNIRCSLQ-----NSRLH 110
DB 74 KMSIDPRVILIVN---FPTQY-----SIMQNDVYDEGPYTCGYQDNHMKTSRVH 122
OY 111 GSAYLTVQVMELEFIPSNLVVAENEPCEVTCL---PSHWTRLPDISWEIGLLVSHSSY 166
DB 123 ---LIVQVPPQIMNISSDIVNEISSVTLCLAIGRPE-----PTVWTR-----HLNV 167
OY 167 Y----FVPEPSDIQSAVSILALTPQSGNTLTCVATWKSILKARKSATVNLTV-----IRCP 217
DB 168 KEGGQFVSEDEYLE-----ISDIKRDQSEYEYCSAL-NDVAAPDVRYKVKITVNPYPYISKA 222
OY 218 QDTGGGINIPGVL---SLPSLIGFSLPTWGK---VGLGLAGTML 255
DB 223 KNTGVSVGQKGIISCASAVPMAEFQ---WFKEDTRLATGLDGIVRI 265

```

Search completed: April 28, 2003, 21:09:17
Job time : 27.0659 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2003, 11:05:56 ; Search time 19.0659 Seconds

(without alignments)
1622.277 Million cell updates/sec

Title: US-09-729-264-6

2077

Perfect score: 1 MERHLITVEAVGSSGSGNEV.....HFOASPNLASPEKVSNTTVV 386

Sequence:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145.5	7.0	405	1 US-08-755-235-4	Sequence 4, Appl1
2	138.5	6.7	2473	9 US-10-184-644-559	Sequence 559, App
3	138.5	6.7	2473	9 US-10-184-634-559	Sequence 559, App
4	134	6.5	708	9 US-10-174-590-584	Sequence 584, App
5	134	6.5	708	9 US-10-176-758-584	Sequence 584, App
6	134	6.5	708	9 US-10-175-727-584	Sequence 584, App
7	134	6.5	708	9 US-10-173-706-584	Sequence 584, App
8	134	6.5	708	9 US-10-175-738-584	Sequence 584, App
9	134	6.5	708	9 US-10-175-738-584	Sequence 584, App
10	134	6.5	708	9 US-10-176-482-584	Sequence 584, App
11	134	6.5	708	9 US-10-176-757-584	Sequence 584, App
12	134	6.5	708	9 US-10-176-913-584	Sequence 584, App
13	134	6.5	708	9 US-10-180-552-584	Sequence 584, App
14	134	6.5	708	9 US-10-180-557-584	Sequence 584, App
15	134	6.5	708	9 US-10-173-700-584	Sequence 584, App
16	134	6.5	708	9 US-10-174-572-584	Sequence 584, App
17	134	6.5	708	9 US-10-174-579-584	Sequence 584, App
18	134	6.5	708	9 US-10-174-582-584	Sequence 584, App
19	134	6.5	708	9 US-10-174-588-584	Sequence 584, App

20	134	6.5	708	9	US-10-175-739-584	Sequence 584, App
21	134	6.5	708	9	US-10-175-740-584	Sequence 584, App
22	134	6.5	708	9	US-10-175-743-584	Sequence 584, App
23	134	6.5	708	9	US-10-176-488-584	Sequence 584, App
24	134	6.5	708	9	US-10-176-492-584	Sequence 584, App
25	134	6.5	708	9	US-10-176-747-584	Sequence 584, App
26	134	6.5	708	9	US-10-176-750-584	Sequence 584, App
27	134	6.5	708	9	US-10-176-985-584	Sequence 584, App
28	134	6.5	708	9	US-10-176-987-584	Sequence 584, App
29	134	6.5	708	9	US-10-176-991-584	Sequence 584, App
30	134	6.5	708	9	US-10-176-992-584	Sequence 584, App
31	134	6.5	708	9	US-10-176-993-584	Sequence 584, App
32	134	6.5	708	9	US-10-184-658-584	Sequence 584, App
33	134	6.5	708	9	US-10-173-695-584	Sequence 584, App
34	134	6.5	708	9	US-10-173-697-584	Sequence 584, App
35	134	6.5	708	9	US-10-173-705-584	Sequence 584, App
36	134	6.5	708	9	US-10-174-576-584	Sequence 584, App
37	134	6.5	708	9	US-10-174-585-584	Sequence 584, App
38	134	6.5	708	9	US-10-174-586-584	Sequence 584, App
39	134	6.5	708	9	US-10-175-747-584	Sequence 584, App
40	134	6.5	708	9	US-10-176-481-584	Sequence 584, App
41	134	6.5	708	9	US-10-176-485-584	Sequence 584, App
42	134	6.5	708	9	US-10-176-487-584	Sequence 584, App
43	134	6.5	708	9	US-10-176-493-584	Sequence 584, App
44	134	6.5	708	9	US-10-176-756-584	Sequence 584, App
45	134	6.5	708	9	US-10-176-911-584	Sequence 584, App

ALIGNMENTS

RESULT 1
US-08-755-235-4
Sequence 4, Application US/08755235
Publication No. US20030059423A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Schmidt, Ann Marie
APPLICANT: Wu, Jun
TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
FILE REFERENCE: 0575/50159
CURRENT APPLICATION NUMBER: US/08/755, 235
CURRENT FILING DATE: 1996-11-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 405
TYPE: PRT
ORGANISM: Human
US-08-755-235-4

Query Match 7.0%; Score 145.5; DB 1; Length 405;
Best Local Similarity 23.5%; Pred. No. 0.0034;
Matches 91; Conservative 38; Mismatches 111; Indels 147; Gaps 20;

QY	9	PEAVSGSGNEVIEGPONATVLKGSQARENCYSGQ---WKLMAALSDMVYLSVPMPE	65
DB	124	PEYDSAS--ELTAGVFN-----KVCTCVSEGYPAGLTSMHLDG-----RP	163
QY	66	ITINDRTS-----ORDOGNFT--SEMIHNVPSDSGNIR-----CSLONRLGSAV	114
DB	164	LVPKEKVSVAKEOTRRPELTGLTLOSELM---VTPARGDPPRTSCSPPLPHRRL	220
QY	115	LTVQVMGEFLP---SVNLVAENP-----CEVTLPSHMTLPDISME	156
DB	221	RTAIDGRWPEPVLEEVQLVV---EPEGAVAPGVTLTCEVPAPPS-----POIHW	272
QY	157	LGLLVSHSYTFVPEPDDSAVSITALTQSNGLTCVATWMSLKARKS--ATVNLTVR	215
DB	273	KD-----GVLPLPPSPVLLTPEIGPDGQTSVATHSHGQESRAVVSISLE	323
QY	216	CPDGTGGINIPGVLSLPSLIGFSLPTWKGKVGGLAGTMLLT-----PTCTLTIRCC	268


```
US-10-176-758-584
; Sequence 584, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-584
```

```
Query Match 6.5%; Score 134; DB 9; Length 708;
Best Local Similarity 23.7%; Pred. No. 0.045;
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;
```

```
QY 14 SSGSNEVIEGPONATVYKSGQARFNCYSGQKLMWALSDMVVLSVRPMEPIITNDRT 73
DB 20 AGSPHFLOQPEDLVLLGEARLPCALGAYGVLVQWTKRSGLAGGQ-----DLPG 71
QY 74 SORYDOGNFTS---EMIIHNVPSDSGNIRCSLONSRLHG-SAYLVV-----QVNGE 122
DB 72 WSRWISGNANQOHDLHIREVELEDEASYECQATQAGLRSPALHVLVPEAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTCPSHWTR-LPDISW-ELGLVSHSSY--FYVE--PSDLQ 176
DB 132 ---PSVSLVA--GVPANLTCRSRGDARPTPELLMFRDGVLLDGATFHQTLKEGTPGSVE 186
QY 177 SAVSIALTPQSNGLTCAVATWKSILKARKSATVNLTVIRCPDPT 220
DB 187 STLTLPFSHDDGATFVCRARSQALPTGRDPAITLSIQYPEPT 230
```

RESULT 6

```
US-10-175-737-584
; Sequence 584, Application US/10175737
; Publication No. US2003003135A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
```

```
LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-584
```

```
Query Match 6.5%; Score 134; DB 9; Length 708;
Best Local Similarity 23.7%; Pred. No. 0.045;
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;
```

```
QY 14 SSGSNEVIEGPONATVYKSGQARFNCYSGQKLMWALSDMVVLSVRPMEPIITNDRT 73
DB 20 AGSPHFLOQPEDLVLLGEARLPCALGAYGVLVQWTKRSGLAGGQ-----DLPG 71
QY 74 SORYDOGNFTS---EMIIHNVPSDSGNIRCSLONSRLHG-SAYLVV-----QVNGE 122
DB 72 WSRWISGNANQOHDLHIREVELEDEASYECQATQAGLRSPALHVLVPEAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTCPSHWTR-LPDISW-ELGLVSHSSY--FYVE--PSDLQ 176
DB 132 ---PSVSLVA--GVPANLTCRSRGDARPTPELLMFRDGVLLDGATFHQTLKEGTPGSVE 186
QY 177 SAVSIALTPQSNGLTCAVATWKSILKARKSATVNLTVIRCPDPT 220
DB 187 STLTLPFSHDDGATFVCRARSQALPTGRDPAITLSIQYPEPT 230
```

RESULT 7

```
US-10-173-706-584
; Sequence 584, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C107
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-584
```

```
Query Match 6.5%; Score 134; DB 9; Length 708;
Best Local Similarity 23.7%; Pred. No. 0.045;
Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;
```

```
QY 14 SSGSNEVIEGPONATVYKSGQARFNCYSGQKLMWALSDMVVLSVRPMEPIITNDRT 73
DB 20 AGSPHFLOQPEDLVLLGEARLPCALGAYGVLVQWTKRSGLAGGQ-----DLPG 71
QY 74 SORYDOGNFTS---EMIIHNVPSDSGNIRCSLONSRLHG-SAYLVV-----QVNGE 122
DB 72 WSRWISGNANQOHDLHIREVELEDEASYECQATQAGLRSPALHVLVPEAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTCPSHWTR-LPDISW-ELGLVSHSSY--FYVE--PSDLQ 176
DB 132 ---PSVSLVA--GVPANLTCRSRGDARPTPELLMFRDGVLLDGATFHQTLKEGTPGSVE 186
QY 177 SAVSIALTPQSNGLTCAVATWKSILKARKSATVNLTVIRCPDPT 220
```

US-10-175-738-584
US-10-175-738-584
; Sequence 584, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION.

Query Match	6.5%;	Score 134;	DB 9;	Length 708;
Best Local Similarity	23.7%;	Pred. No. 0.045;		
Matches 53; Conservative	40;	Mismatches 101;	Indels 30;	Gaps 10

RESULT 9
US-10-175-752-584

Publication No. US20030022295A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C60
CURRENT APPLICATION NUMBER: US/10/175,752

Query Match	6.5%;	Score 134;	DB 9;	Length 708;
Best Local Similarity	23.7%;	Pred. No. 0.045;		
Matches	53;	Conservative	40;	Mismatches 101; Indels 30; Gaps 10;

RESULT 10
US-10-176-482-584
; Sequence 584, Application US/10176482

Query Match	6.5%;	Score 134;	DB 9;	Length 708;
Best Local Similarity	23.7%;	Pred. No. 0.045;		
Matches	53;	Conservative	40;	Mismatches 101;
			Indels	30;
			Gaps	10

	Matches	53: Conservative	40: Mismatches	101: Indels	30: Caps
QY	14	SSGSNEVIEPONTATLKSGARPNCTVSGKMLIMMALSDMYLVLSVRPMETITNDRET	73		
Db	20	APPSPHFIQGPBDLVYLIGEARAPCALGAYWLVMTKSLATLGGQR-----DLPG	71		
QY	74	SGRYDGGNGFTS---EMLIHNHVPSSDGNRCISLQNSRLHG--SAVLTY-----QVNGE	122		
Db	72	WRYWYISGNMAANGDHLHPIRVELEDEASVECATGAGLRSPADLHVLPPEARQYVLG	133		
QY	123	LFIPSVNLVYANENPECVETCLPSHWTR-LPDISW-ELGLIVSHSSTY--FYVE--PSDLO	176		

Db 132 ---PSVSIVA--GVPANLTCRSRGDARPTPELLMFRCGVLLDGAFTHQTLKECTPGSVE 186
 QY 177 SAVSIIALTPOSGNGLTCVATWTKSLKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVCRARSQALPTGRDTAITLSLQYRPEVT 230

RESULT 11

US-10-176-757-584
 ; Sequence 584, Application US/10176757
 ; Publication No. US20030022297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C86
 CURRENT APPLICATION NUMBER: US/10/176,757
 CURRENT FILING DATE: 2002-06-20
 Prior Application removed - See file wrapper or Palm

SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-176-757-584

Query Match 6.5%; Score 134; DB 9; Length 708;
 Best Local Similarity 23.7%; Pred. No. 0.045;
 Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY 14 SSGSNEVEIEGPONATVTKGSOARFNCTVSQGWKLIMALSDMVLVSVPMEPIITNDRET 73
 Db 20 AGSPHFLQOPEDLVLLGEARLPCALGAYWGLVQWTKSGIALGGR-----DLPG 71
 QY 74 SQRVDGNGFTS---EMIIHNVEPSDSGNIRCSLQNSRLHG-SAYLTV-----QVNGE 122
 Db 72 WSRWISGNANANGQHDHIREVELEDEASTYEQATQAGLSRPAQLHVLVPEAPQVLLG 131
 QY 123 LFIISVNLVAENEPCEVTCIPSHWTR-LPDISW-ELGLVSHSY--FVPE--PSDLQ 176
 Db 132 ---PSVSIVA--GVPANLTCRSRGDARPTPELLMFRCGVLLDGAFTHQTLKECTPGSVE 186
 QY 177 SAVSIIALTPOSGNGLTCVATWTKSLKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVCRARSQALPTGRDTAITLSLQYRPEVT 230

RESULT 12

US-10-176-913-584
 ; Sequence 584, Application US/10176913
 ; Publication No. US20030022298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C66
 CURRENT APPLICATION NUMBER: US/10/176,913
 CURRENT FILING DATE: 2002-06-20
 Prior Application removed - See file wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-176-913-584

Query Match 6.5%; Score 134; DB 9; Length 708;
 Best Local Similarity 23.7%; Pred. No. 0.045;
 Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY 14 SSGSNEVEIEGPONATVTKGSOARFNCTVSQGWKLIMALSDMVLVSVPMEPIITNDRET 73
 Db 20 AGSPHFLQOPEDLVLLGEARLPCALGAYWGLVQWTKSGIALGGR-----DLPG 71
 QY 74 SQRVDGNGFTS---EMIIHNVEPSDSGNIRCSLQNSRLHG-SAYLTV-----QVNGE 122
 Db 72 WSRWISGNANANGQHDHIREVELEDEASTYEQATQAGLSRPAQLHVLVPEAPQVLLG 131
 QY 123 LFIISVNLVAENEPCEVTCIPSHWTR-LPDISW-ELGLVSHSY--FVPE--PSDLQ 176
 Db 132 ---PSVSIVA--GVPANLTCRSRGDARPTPELLMFRCGVLLDGAFTHQTLKECTPGSVE 186
 QY 177 SAVSIIALTPOSGNGLTCVATWTKSLKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVCRARSQALPTGRDTAITLSLQYRPEVT 230

RESULT 13

US-10-180-552-584
 ; Sequence 584, Application US/10180552
 ; Publication No. US20030022300A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C153
 CURRENT APPLICATION NUMBER: US/10/180,552
 CURRENT FILING DATE: 2002-06-25
 Prior Application removed - See file wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-180-552-584

Query Match 6.5%; Score 134; DB 9; Length 708;
 Best Local Similarity 23.7%; Pred. No. 0.045;
 Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY 14 SSGSNEVEIEGPONATVTKGSOARFNCTVSQGWKLIMALSDMVLVSVPMEPIITNDRET 73
 Db 20 AGSPHFLQOPEDLVLLGEARLPCALGAYWGLVQWTKSGIALGGR-----DLPG 71
 QY 74 SQRVDGNGFTS---EMIIHNVEPSDSGNIRCSLQNSRLHG-SAYLTV-----QVNGE 122

Db 72 WSRWISGNMANGOHDLHTRPELEDEASYECAQTQAGLRSPADLHVLVPEAPQVLGG 131
 QY 123 LFIPIVNLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLVSHSSY--FVPE--PSDLQ 176
 Db 132 ---PSVSLVA--GVANLTCRSRGDARPTPELLMFRDGLLDGATFHQTLKEGTGSVE 186
 QY 177 SAVSIALTPQSNGLTCVATWKSLSKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVFCARASQALPTGRDTAITLSIQYPPVPT 230

RESULT 14

US-10-180-557-584
 ; Sequence 584, Application US/10180557
 ; Publication No. US2003002201A1
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3430R1C14
 CURRENT APPLICATION NUMBER: US/10/180,557
 CURRENT FILING DATE: 2002-06-25
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-180-557-584

Query Match 6.5%; Score 134; DB 9; Length 708;
 Best Local Similarity 23.7%; Pred. No. 0.045;
 Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY 14 SSGNEVIEGPNATVLSKQARFNCTVSQMKLIMWALSDMVLSVRPEPIITNDRET 73
 Db 20 AGSPPHLQPEDLVLLGEBARLPALGATWGIWQTKSGIALGGR-----DLPG 71
 QY 74 SQRYDGGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG--SAYLTV-----QVNGE 122
 Db 72 WSRWISGNMANGOHDLHTRPELEDEASYECOATQAGLRSPADLHVLVPEAPQVLGG 131
 QY 123 LFIPIVNLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLVSHSSY--FVPE--PSDLQ 176
 Db 132 ---PSVSLVA--GVANLTCRSRGDARPTPELLMFRDGLLDGATFHQTLKEGTGSVE 186
 QY 177 SAVSIALTPQSNGLTCVATWKSLSKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVFCARASQALPTGRDTAITLSIQYPPVPT 230

RESULT 15

US-10-173-700-584
 ; Sequence 584, Application US/10173700
 ; Publication No. US20030027262A1
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James

APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3430R1C14
 CURRENT APPLICATION NUMBER: US/10/173,700
 CURRENT FILING DATE: 2002-06-17
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-173-700-584

Query Match 6.5%; Score 134; DB 9; Length 708;
 Best Local Similarity 23.7%; Pred. No. 0.045;
 Matches 53; Conservative 40; Mismatches 101; Indels 30; Gaps 10;

QY 14 SSGNEVIEGPNATVLSKQARFNCTVSQMKLIMWALSDMVLSVRPEPIITNDRET 73
 Db 20 AGSPPHLQPEDLVLLGEBARLPALGATWGIWQTKSGIALGGR-----DLPG 71
 QY 74 SQRYDGGNFTS---EMIHNVPSDSGNIRCSLQNSRLHG--SAYLTV-----QVNGE 122
 Db 72 WSRWISGNMANGOHDLHTRPELEDEASYECAQTQAGLRSPADLHVLVPEAPQVLGG 131
 QY 123 LFIPIVNLVAENEPCEVTCPLPSHWTR-LPDISW-ELGLVSHSSY--FVPE--PSDLQ 176
 Db 132 ---PSVSLVA--GVANLTCRSRGDARPTPELLMFRDGLLDGATFHQTLKEGTGSVE 186
 QY 177 SAVSIALTPQSNGLTCVATWKSLSKARKSATVNLTVIRCPDPT 220
 Db 187 STLTLPFSHDDGATFVFCARASQALPTGRDTAITLSIQYPPVPT 230

Search completed: April 29, 2003, 20:17:34
 Job time : 27.0659 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 18:14:43 ; Search time 14.0485 seconds

(without alignments)
808.430 Million cell updates/sec

Title: US-09-729-264-6

Perfect score: 2077

Sequence: 1 MERHLTVEAVGSGSGNEV.....HPQASFNLASPEKVSNTTVV 386

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	128.5	6.2	1345	2	US-08-977-767-3
2	128	6.2	1447	4	US-09-041-886-25
3	128	6.2	1447	5	PCT-US94-05277-2
4	127.5	6.1	869	1	US-08-374-834-16
5	127.5	6.1	869	2	US-08-644-271-16
6	127.5	6.1	869	4	US-09-077-955-33
7	127	6.1	332	4	US-09-062-365-1
8	124.5	6.0	340	4	US-09-651-200-2
9	124.5	6.0	441	4	US-09-651-200-4
10	123.5	5.9	534	4	US-09-651-200-6
11	123.5	5.9	534	4	US-09-651-200-24
12	122	5.9	318	2	US-08-633-148-4
13	122	5.9	340	2	US-08-633-148-2
14	114	5.5	868	1	US-08-374-834-1
15	114	5.5	868	2	US-08-644-271-1
16	114	5.5	868	4	US-09-077-955-1
17	114	5.5	1395	4	US-09-540-245A-15
18	113	5.4	689	4	US-09-499-964-1
19	113	5.3	365	2	US-08-979-424-3
20	111	5.3	365	4	US-09-272-496-2
21	109.5	5.3	325	4	US-09-651-200-20
22	109.5	5.3	478	5	PCT-US95-08493-15
23	109.5	5.3	860	5	PCT-US95-08493-19
24	109.5	5.3	868	5	PCT-US95-08493-21
25	109.5	5.3	890	1	US-08-445-640-2
26	109.5	5.3	890	3	US-08-170-558-2
27	109.5	5.3	890	3	US-08-447-314-2

28	109.5	5.3	890	3	US-08-445-641-2	Sequence 2, Appli
29	109.5	5.3	911	1	US-08-286-305A-1	Sequence 1, Appli
30	109.5	5.3	911	2	US-08-441-104A-1	Sequence 1, Appli
31	109.5	5.3	911	2	US-08-440-816A-1	Sequence 1, Appli
32	109.5	5.3	911	4	US-09-417-381A-1	Sequence 1, Appli
33	107.5	5.2	362	1	US-08-415-751-6	Sequence 6, Appli
34	106.5	5.1	946	5	PCT-US95-08493-13	Sequence 13, Appli
35	106	5.1	801	1	US-07-906-349A-6	Sequence 6, Appli
36	105	5.1	319	1	US-08-597-495B-22	Sequence 22, Appli
37	105	5.1	319	4	US-09-068-051A-22	Sequence 22, Appli
38	105	5.1	319	4	US-09-336-536-67	Sequence 67, Appli
39	105	5.1	319	4	US-09-254-465A-6	Sequence 6, Appli
40	105	5.1	365	4	US-08-928-383B-2	Sequence 2, Appli
41	105	5.1	690	4	US-08-935-433-2	Sequence 2, Appli
42	105	5.1	690	4	US-09-553-132-2	Sequence 17, Appli
43	105	5.1	1297	4	US-09-540-245A-17	Sequence 16, Appli
44	105	5.1	1381	4	US-09-540-245A-16	Sequence 16, Appli
45	104.5	5.0	879	1	US-08-554-612C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-977-767-3
Sequence 3, Application US/08977767

Patent No. 5972684

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Yue, Henry

APPLICANT: Greenwald, Sara

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: CARBONIC ANHYDRASE VIII

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/977,767

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

FILING DATE:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0423 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1345 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 1532042

US-08-977-767-3

Query Match

6.2%; Score 128.5; DB 2; Length 1345;

Best Local Similarity 33.0%; Pred. No. 0.0098;
Matches 37; Conservative 1; Mismatches 41; Indels 33; Gaps 5;

QY 190 GTLTCVATWKSILKARKSATVNLVIRCPDGTGGI-----NIPGVLSLPSLPSLPTWCK 245
DB 414 GTCTCTGT-----GC-GCTGGAAAGCCTCAGAGCCCGCTGATGTGA 455

QY 246 VGLGLAGTMLLT-PICTLTRCCCRRCGCCGCCRC-----CFCC 286
DB 456 CGTGAAGAGCGTCTCTATGACCCCTTCTGCCCTGAGACACAGACACC 507

RESULT 2

US-09-041-886-25
Sequence 25, Application US/09041886
Patent No. 6235872

GENERAL INFORMATION:

APPLICANT: Bredesen, Dale E.
APPLICANT: Rabizadeh, Sharoz
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041.886
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-886-25

Query Match 6.2%; Score 128; DB 4; Length 1447;
Best Local Similarity 24.8%; Pred. No. 0.012;
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SSGSNEV-----IEGPNATVYKSGQARFNCYVSGW--KLIMWLSDM 55
DB 220 SRTGNEAEVRIISDPGLHROLVFLQRPNSVVAIEGKDAVECCVS-GYPPSPFTLNGEE 278
QY 56 VLVSRPMEPIITNDRFTSORYDOGNFTSEMIITHNVEPSDSGNIRC--SLONSRLHGA 113
DB 279 VI-----QLRSKYSILGG--SNLLISNVTDDSGMTVCVYTKNENISASA 323
QY 114 YLVQVOMGELFIPSVNLVAENEPCEVTCPLPSHWRLPDISW-ELGLVSHSSYFVEPP 172
DB 324 ELTVLVPPEFNLHPNSLVAYESMDIEFECTVS-GKVPPTVMNMKNQDVVIPSDFQIVGG 382
QY 173 SDLOSASVIALTPQSNGLTCVATWKSILKARKSATVNLVIRCPDGTGGGGINIGVLS 232
DB 383 SNLR-----ILGVKSDSEGTYOCVAENAGNAOTSQOLIVKPAIPSSS-----VLPS 430

QY 233 LP 234
DB 431 AP 432

RESULT 3

PCT-US94-05277-2
Sequence 2, Application PC/TU9405277

GENERAL INFORMATION:

APPLICANT: Bruskin, Arthur
APPLICANT: Jarosz, David E.
APPLICANT: Johnson, Karen
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
APPLICANT: Zaretsky, James R.
TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05277
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42709
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05277-2

Query Match 6.2%; Score 128; DB 5; Length 1447;
Best Local Similarity 24.8%; Pred. No. 0.012;
Matches 60; Conservative 33; Mismatches 99; Indels 50; Gaps 10;

QY 14 SSGSNEV-----IEGPNATVYKSGQARFNCYVSGW--KLIMWLSDM 55
DB 220 SRTGNEAEVRIISDPGLHROLVFLQRPNSVVAIEGKDAVECCVS-GYPPSPFTLNGEE 278
QY 56 VLVSRPMEPIITNDRFTSORYDOGNFTSEMIITHNVEPSDSGNIRC--SLONSRLHGA 113
DB 279 VI-----QLRSKYSILGG--SNLLISNVTDDSGMTVCVYTKNENISASA 323
QY 114 YLVQVOMGELFIPSVNLVAENEPCEVTCPLPSHWRLPDISW-ELGLVSHSSYFVEPP 172
DB 324 ELTVLVPPEFNLHPNSLVAYESMDIEFECTVS-GKVPPTVMNMKNQDVVIPSDFQIVGG 382
QY 173 SDLOSASVIALTPQSNGLTCVATWKSILKARKSATVNLVIRCPDGTGGGGINIGVLS 232
DB 383 SNLR-----ILGVKSDSEGTYOCVAENAGNAOTSQOLIVKPAIPSSS-----VLPS 430
QY 233 LP 234
DB 431 AP 432

QY 278 -----CCCRCC-----FCCRRRGRFRIOPQKSEKETEETESGNGNSGNSDEQKT 327
 Db 389 TPIPIREYCLAVKEFLCAKE-----WLVMEKTHRGILYSEMHLLSYECSKLPJMHMDP 444
 QY 328 TETASIPRPSCESSDEQONSSCGPPHQAADRP-----PRASHQAASFNLA 375
 Db 445 TACARLP-----HLDYKKNELKTFPP--MTSSKPSVDIPNLPSSSSSSFSVS 489

RESULT 6

US-09-077-955-33
 Sequence 33, Application US/09077955A
 Patent No. 6413740

GENERAL INFORMATION:
 APPLICANT: Valenzuela et al., David M.
 TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
 FILE REFERENCE: REG195-B-PCT-US
 CURRENT APPLICATION NUMBER: US/09/077,955A
 EARLIER FILING DATE: 1998-09-10
 EARLIER APPLICATION NUMBER: PCT/US96/20696
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 08/644,271
 EARLIER FILING DATE: 1996-05-10
 EARLIER APPLICATION NUMBER: 60/008,657
 EARLIER FILING DATE: 1995-12-15
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 33
 LENGTH: 869
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-077-955-33

Query Match 6.1%; Score 127.5; DB 4; Length 869;
 Best Local Similarity 20.9%; Pred. No. 0.0065;
 Matches 86; Conservative 51; Mismatches 176; Indels 99; Gaps 19;

QY 19 EYIEGPONATYVKGSOARFNCTVSQGMK-LIMMALSDMVLSVPRMEPIITNDRFTSQRY 77
 Db 122 KITRPPINVKIIEGLKAVLPCTMGNPKRPSVMIKGD-----SPLRENSRIAYLE- 171
 QY 78 DQGNMTSEMITIHVNEPDSGNIRCSLQNSRLHGSAY-LTVQVNGELFIPRVNVLVAENE 136
 Db 172 -----SSLRHYNQKEDAGQRCYVAKNSL--GTAISKVVKLEVEFARILRAPESINLV 223
 QY 137 P-----CEVTLPSHMTLRLDISW-ELGLVSHSSYFVPPSPDLOSASVILATLPQ 187
 Db 224 TFGSFVTLHCTATGIP-----VPIITWLENGNAVSSGSIQESVQDRAVIDSLQLFITKP- 277
 QY 188 SNGTLTCVAT-----WKSIAKRSATVNLTVIRCPQDIDGG-----INIPGVLSLP 234
 Db 278 --GLYTCTATNKHGEKSTAAKAAATISIAEMSKPKQDNKKGCAQYKEVCNAVLAKDALV 335
 QY 235 SLGFSI-----PTWKGVLGLAGTMLTPTCTLIRCCCRRCRCGCGN----- 277
 Db 336 FLNITSYADPEEAQELVHTANL-----KVSVPICRAAEALLCNHIFQECSPGVVP 388
 QY 278 -----CCCRCC-----FCRRKRKRFRIQFOKSKSEKTEKTEESGNGNSGNSDEQKT 327
 Db 389 TPIPIREYCLAVKEFLCAKE-----WLVMEKTHRGILYSEMHLLSYECSKLPJMHMDP 444
 QY 328 TETASIPRPSCESSDEQONSSCGPPHQAADRP-----PRASHQAASFNLA 375
 Db 445 TACARLP-----HLDYKKNELKTFPP--MTSSKPSVDIPNLPSSSSSSFSVS 489

RESULT 7

US-09-062-365-1
 Sequence 1, Application US/09062365
 Patent No. 6463422

GENERAL INFORMATION:
 APPLICANT: Schmidt, Ann Marie
 APPLICANT: Stern, David

TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
 TITLE OF INVENTION: SUBJECT
 FILE REFERENCE: 55424
 CURRENT APPLICATION NUMBER: US/09/062,365
 CURRENT FILING DATE: 1998-04-17
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 1
 LENGTH: 332
 TYPE: PRT
 ORGANISM: Human
 US-09-062-365-1

Query Match

6.1%; Score 127; DB 4; Length 332;
 Best Local Similarity 25.0%; Pred. No. 0.0019;
 Matches 70; Conservative 31; Mismatches 87; Indels 92; Gaps 15;

QY 9 PEAVSGSGNEVIEGPONATVKGSOARFNCTVSQGMK-LIMMALSDMVLSVPRMEPI 65
 Db 102 PEYDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTISWHLDG-----KP 141
 QY 66 IITNDRFTS-----QRDQCGNT--SEMITIHVNEPDSGNIR-----CSLQNSRLHGSAY 114
 Db 142 LVPRKGVSVKEQTRRHETGELTLOELM--VTPARAGDPRPTSCSFSFGLPRHRAL 198
 QY 115 LTVQVNGELFIP-----SVNLVAENEP-----CEVTLPSHMTLRLDISW 156
 Db 199 RTAPIQPRVMEPPVLEVLQV--EPEGAVAPGGCTVITLCEVPAPPS-----PQIHWM 250
 QY 157 LGLVSHSSYFVPPSPDLOSASVILATLPQNSGTLTCVATWKSIAKRSATVNLTVIRC 216
 Db 251 KD-----GVPLPLPSPVLLILPEIGPDQGTYSVATSHSGPQSRVAVSIIIE- 300
 QY 217 PDITGGGINIPGVLSLPSLGFSLPTWKGVLGLACTMLL 256
 Db 301 PGEKG-----PTAGSVGSGGLCTLAL 321

RESULT 8

US-09-651-200-2
 Sequence 2, Application US/09651200
 Patent No. 6429303

GENERAL INFORMATION:
 APPLICANT: Green et al
 TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 TITLE OF INVENTION: Polypeptides Encoded Thereby
 FILE REFERENCE: 15966-562 (CURA-62)
 CURRENT APPLICATION NUMBER: US/09/651,200
 CURRENT FILING DATE: 2000-08-30
 PRIOR APPLICATION NUMBER: 60/152383
 PRIOR FILING DATE: 1999-09-03
 PRIOR APPLICATION NUMBER: 60/172909
 PRIOR FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: 60/183578
 PRIOR FILING DATE: 2000-02-18
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 2
 LENGTH: 340
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-651-200-2

Query Match 6.0%; Score 124.5; DB 4; Length 340;
 Best Local Similarity 21.2%; Pred. No. 0.0033;
 Matches 77; Conservative 51; Mismatches 127; Indels 109; Gaps 17;

QY 14 SSGNEVIEGPONATV-LKSGARFNCTVS--QGMKL-----IMMALSDMVLSVPRMEPI 66
 Db 48 SPTGAVEGVPEPDVVALVGTATLHCSFSPERGSLVQNLIMQDLITPTKQLV----- 100
 QY 67 ITNDRFTSQRYDQGNF-----TSEMITIHVNEPDSGNIRCSLQNSRLHGS 112


```

: GENERAL INFORMATION:
: APPLICANT: Green et al
: TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
: TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
: TITLE OF INVENTION: Polypeptides Encoded Thereby
: FILE REFERENCE: 15966-562 (CURA-62)
: CURRENT FILING DATE: 2000-08-30
: PRIOR APPLICATION NUMBER: US/09/651,200
: PRIOR FILING DATE: 1999-09-03
: PRIOR APPLICATION NUMBER: 60/172909
: PRIOR FILING DATE: 1999-12-21
: PRIOR APPLICATION NUMBER: 60/183578
: PRIOR FILING DATE: 2000-02-18
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 24
: LENGTH: 534
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Sequence
: OTHER INFORMATION: m25020.protein from Figure 4.
US-09-651-200-24

```

```

Query Match          5.9%; Score 123.5; DB 4; Length 534;
Best Local Similarity 21.2%; Pred. No. 0.0076;
Matches 77; Conservative 50; Mismatches 128; Indels 109; Gaps 17;

```

```

QY 14 SGGGNEVIEGPNATV-LKGSQARFNCTVSG--QMKL---IMMALSDMNVLSVPRMEPI 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 SPGAVEVYVPPEDPVVALVGTATLRCSPSPGFSLAOLNIMWLTDTKOLV----- 294
QY 67 ITNDRFSQRYDQGNF-----TSEMIHNVPSDSGNIRCSLQNSRLHGS 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 295 ---HSTTBGR-DGSAIYAKRTALFPDLAOGNASLRLOQVRVADSGSFTCFV-STRDGGS 349
QY 113 AYLTVOMGELFIPSVNLV---VAENEPCEVTCPLPSHTRLP--DISNELGL---LVSH 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 AAVSLQVAPYSKPSMTLPBNKDLRPGDTVTITC---SSYRGYPEAEVFMQDQGVPLNGN 407
QY 164 SSYTFPEPSDDQSAVSIILALTPQSNGLTCVATKSLKARSAVNLTVIRC--QDT 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 VITSQANQGLFVHSVLRVVLGANGTYS-----LVNRPVLAQDDA 449
QY 221 GGINIPGVLSLPSLGSFLPTMGKVGGLAGTMLTPTCTLTIRCCCRRCGCCNCCC 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 450 HGSVITTGQPMTFPPAL---WTVYGLSVCLLALLV----- 482
QY 281 RCFCCRRRRGRFRIQFKKSEKTKNKETETESGNENSGYNSDEQKTETASLPKSCGS 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 ALAFVCMRK-----IKQSCFEENAGAEQDDG-----EGGSKTALQPLKHSDS 525
QY 341 SDPE 344
DB 526 KEDD 529

```

```

RESULT 12
US-08-633-148-4
: Sequence 4, Application US/08633148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSEY, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO

```

```

: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/633,148
: FILING DATE: 16-APR-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY ESO., MATTHEW B.
: REGISTRATION NUMBER: 39,787
: REFERENCE/DOCKET NUMBER: 014618-005600US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 318 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-633-148-4

```

```

Query Match          5.9%; Score 122; DB 2; Length 318;
Best Local Similarity 24.9%; Pred. No. 0.005;
Matches 69; Conservative 30; Mismatches 86; Indels 92; Gaps 15;

```

```

QY 9 PEAVSGSGNEVIEGPNATV-LKGSQARFNCTVSG--WKLIMMALSDMNVLSVPRMEPI 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102 PEYDSAS--ELTAGVFN-----KVGTCVSESGYPAGTILSMHLDG-----KP 141
QY 66 ITNDRFTS---QRDQGNFT---SEMIHNVPSDSGNIR---CSLQNSRLHGSAY 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 142 LVNPEKGVSEYQRTRHETGLFTLOSELM---VTPARGDPRPTFSFSGLPFRHRL 198
QY 115 LTVOMGELFIP---SVNLVAENEP-----CEVTCPLPSHTRLPDISWE 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 199 RTAPLPQPMWEPVPLEEYQLV---EPGGAVAPGTVTLTCEVPAOPS-----PQIHM 250
QY 157 LGLVSHSSSYTFPEPSDDQSAVSIILALTPQSNGLTCVATKSLKARSAVNLTVIRC 216
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 KD-----GVPLPLPSPVLILPEIGPDQGTYSQVATHSHGHPQESRAVSISIE 300
QY 217 PQTGGGINIPGVLSLPSLGSFLPTMGKVGGLAGT 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 PGEER-----PTAGSVGGSGLG 318

```

```

RESULT 13
US-08-633-148-2
: Sequence 2, Application US/08633148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSEY, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,148
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY ESO., MATTHEW B.
; REGISTRATION NUMBER: 39,787
; REFERENCE/DOCKET NUMBER: 014618-005600US.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-633-148-2

```

```

Query Match          5.9%; Score 122; DB 2; Length 340;
Best Local Similarity 24.9%; Pred. No. 0.0055;
Matches 69; Conservative 30; Mismatches 86; Indels 92; Gaps 15;

```

```

QY 9 PEAVSGSSGNEVEYEGPQNAVTKGSOARFNCYSOG---WKLIMALSDMVVLSVPMEP 65
DB 124 PELVSDS--ELTAGVFN-----KVGTSSEGSYPAGTSLSHLDG-----RP 163
QY 66 IITNDRETS-----QRYDQGNFT--SEMIHNVEPSDSGNIR---CSLONSRLHGSAY 114
DB 164 LVNKEGVSVKEQTRRHPETGLTLOELM---VTPARGDDPRPTSCSPCLPFRHRL 220
QY 115 LVVQNGELFIP---SVNLVAENED-----CEVTLPSHWTLPDISME 156
DB 221 RTAPIOPRWEPEVPLEVOVLV---EPEGAVAPGCTVTLTCEVPAPS-----POIHMM 272
QY 157 IGLIVSSSYFPEPSDLSAVSILALTPQSNCTLTCAVTKSLKARKSATVNLVIRG 216
DB 273 KD-----GVPILPPSPVLILPEIGPDGTYSCVATHSSHPQESRAVSIITE- 322
QY 217 PDDTGGGINIPGVLSLPSLGSFLPTMGKVGGLAGT 253
DB 323 PGEEG-----PTAGSYGSGSLGT 340

```

```

RESULT 14
US-08-374-834-1
; Sequence 1, Application US/08374834
; Patent No. 5656473
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,834
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/095,658
; FILING DATE: 21-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Covert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 190A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-374-834-1

```

```

Query Match          5.5%; Score 114; DB 1; Length 868;
Best Local Similarity 21.3%; Pred. No. 0.1;
Matches 80; Conservative 40; Mismatches 150; Indels 106; Gaps 18;

```

```

QY 19 EVIEGPQNAVTKGSOARFNCYSOGWK-LIMALSDMVVLSVPMEPITNDRETSQR 77
DB 122 KITRPPIVKKIIEGLKAVLPCTTMGNPKPVSWMKGDALRE-----NSRIAVLE- 171
QY 78 DCGNFTSEMIHNVEPSDSGNIRCSLONSRLHGSAY-LTVQNGELFIPSVNLVAENE 136
DB 172 -----SGSLRIHNQKEDAGYRCVAKNSL--GTAISKYKLEVEYARILRAPESINV 223
QY 137 P-----CEVTLPSHWTLPDISW-ELGLVSSHSSYFPEPSDLSAVSIALTPQ 187
DB 224 TFGSEVTLRCPAIGMP-----VPTISWENGVNAGSSGSIQENVKDRVIDSLQFLTRP- 277
QY 188 SNGTLCAVAT---WKSIAKRSATVNLTVIR-----CPDDTGGGINIPGVLSLP 234
DB 278 -GLYTCLATNKGKGFSTAKAAATVSAIEMWSKQESKGYCAOYRGVCAVILVKOSLV 335
QY 235 SLGFSLP-----TWGKVLGLAGTMLLPPTCTLTIRCCCRRCGCGN----- 277
DB 336 FNTSTYPPDEEAOELLITANML-----KAVSPLCRPAEALICHLQESPGVLP 388
QY 278 -----CCCRCC-----FCCRRKRGFRIOFKKSEKTKNETKTESGENSGYNSDEQKT 327
DB 389 TEMPLCREYCLAVKELFCA-----KEMILAMGKTRHGLYRSGMHE--LPV 431
QY 328 TETASLP-----PKSC 338
DB 432 PECSKIPSMHODPTAC 447

```

```

RESULT 15
US-08-644-271-1
; Sequence 1, Application US/08644271
; Patent No. 5814478
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:

```


GenCore version 5.1.4_p5_4578
(c) 1993 - 2003 CompuGen Ltd.

protein search, using sw model

April 28, 2003, 16:24:31 ; Search time 38.4662 Seconds

1337.141 Million cell updates/sec

US-09-729-264-6
2077

1 MERHLLTPEAVGSGSCNEY.....HPQASFNLASPEKVSNTTV 386

BLSUM62	
Gabor 10.0	Ganetx 0.5

908470 seqs, 133250620 residues

hits satisfying chosen parameters: 908470

length: 0
length: 2000000000

Minimum Match 0%

Listing first 45 summaries

A_Geneseq_101002:*

2	/SID52/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2	/SID52/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3	/SID52/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4	/SID52/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5	/SID52/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6	/SID52/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7	/SID52/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8	/SID52/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9	/SID52/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10	/SID52/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11	/SID52/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12	/SID52/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13	/SID52/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14	/SID52/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15	/SID52/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16	/SID52/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17	/SID52/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18	/SID52/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19	/SID52/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20	/SID52/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21	/SID52/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22	/SID52/gcgdata/genseq/genseqp-emb1/AA2001.DAT *
23	/SID52/gcgdata/genseq/genseqp-emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2077	100.0	386	23	AAU75542	Human B7-1ike prot
2	2014	97.0	382	23	AAU75540	Human B7-1ike prot
3	2000	96.3	386	23	AAU75543	Human B7-1ike prot
4	1514	72.9	377	23	AAU75541	Human B7-1ike prot
5	1326	63.8	463	22	ABG28169	Novel human diagno
6	928.5	44.7	370	23	AAU75544	Mouse B7-1ike prot
7	587.5	28.3	631	23	AAU75547	Rat B7-1ike protei
8	579.5	27.9	270	23	AAU75545	Mouse B7-1ike prot
9	498	24.0	223	23	AAU75546	Mouse B7-1ike prot
10	151	7.3	404	22	AAB81925	Extracellular cirt

11	151	7.3	404	23	AAE23219	Human receptor for
12	151	7.3	404	23	AAU77543	Human receptor for
13	151	7.3	404	23	AA48745	Human RAGE protein
14	135	6.5	1496	20	AAAB8030	Melanoma associated
15	135	6.5	1496	21	AAI70469	Human p53 target m
16	134	6.5	1498	22	ABBI1587	Human peroxidasin
17	134	6.5	1592	22	ABBI1587	Human immunoglobulin
18	134	6.5	594	23	ABBI5753	Human pancreas GPR
19	134	6.5	708	22	AAU29315	Human p35, putativ
20	134	6.5	708	22	ABBI5752	Human polypeptide
21	133.5	6.4	344	22	AAAB8713	Deleted in colorec
22	128	6.2	1447	16	AAAB8553	Human DCC protein,
23	128	6.2	1447	20	AAI33496	Human UNC-40 prote
24	128	6.2	1447	22	AAAB5093	Deleted in Colorec
25	128	6.2	1728	12	AAAR1314	Protein MYLK diff
26	128	6.2	1953	23	AAU84351	Alternatively splic
27	127.5	6.1	457	17	AAAB2575	Nsk2 extracellulan
28	127.5	6.1	475	17	AAAP4982	Alternatively splic
29	127.5	6.1	863	19	AAAB2565	Alternatively splic
30	127.5	6.1	867	19	AAAB2583	Mouse receptor tyro
31	127.5	6.1	869	18	AAAB2611	Human muscle-speci
32	127.5	6.1	869	17	AAAB2606	Human Dmk receptor
33	127.5	6.1	871	17	AAAB4087	Nsk2 receptor. M
34	127.5	6.1	871	19	AAAB4052	Mouse receptor tyro
35	127.5	6.1	873	17	AAAB4298	Nsk2 receptor with
36	127.5	6.1	873	19	AAAB2573	Alternatively splic
37	127.5	6.1	881	17	AAAB4091	Alternatively splic
38	127.5	6.1	881	19	AAAB2572	Nsk2 receptor with
39	126	6.1	332	21	AAAY2130	Mouse Nsk2 (altern
40	125.5	6.0	537	22	AAAG7857	Human Receptor to
41	125.5	6.0	576	22	AAAG7856	Mutant protein mmu
42	125.5	6.0	586	22	AAAG7856	Mutant protein mmu
43	125.5	6.0	869	22	AAAB6841	Protein of muscle
44	125	6.0	869	22	AAAB6842	Amino acid sequenc
45	125	6.0	1070	18	AAAB8727	Human colon carcin
			1070	23	ABG61942	Prostate cancer s

ALIGNMENTS

ID	AAU75542	AAU75542 standard; Protein; 386 AA.
AC	AAU75542;	
DT	23-APR-2002 (first entry)	
DE	Human B7-like protein, B7-L_h3.	
FW	Human: B7-like protein; B7-L; antifertility; gynaecological; antitumour; cytostatic; immunosuppressive; antiarthritis; antirheumatic; antinflammatory; dermatological; antipsoriatic; neuroprotective; antidiabetic; haemostatic; antithyroid; antiinicer; antiallergic; antiasthmatic; nephrotoxic; antibacterial; virucide; tumour; cancer; reproductive disorder; graft versus host disease; autoimmune disease; toxic shock syndrome; allergy; nephropathy; skin disorder; endocrinopathy; lymphoproliferative disorder.	
OS	Homo sapiens.	
PN	WO200200710-A2.	
PD	03-JAN-2002.	
PF	28-JUN-2001; 2001WO-US20719.	
PR	28-JUN-2000; 2000US-214512P.	
PR	28-NOV-2000; 2000US-0729264.	
PA	(AMGE-) AMGEN INC.	
WT	Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;	

XX WPI: 2002-130881/17.
 DR N-PSDB: ABK13030.
 XX
 PT New B7-Like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 13; Fig 3; 135pp; English.

XX The invention relates to an isolated B7-Like (B7-L) polypeptide (1).
 CC The polypeptide, polynucleotide encoding it and antibody against (1) are
 CC useful for treating B7-Like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (1) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (1) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allersensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{h3}.

XX Sequence 386 AA:

Query Match 100.0%; Score 2077; DB 23; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1,9e-164;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHHLTVEAVSGSGNVEIEGPONATVYKGSQARFNTVSGMKLIMALSDMWVLSTV 60
 DB 1 MEHHLTVEAVSGSGNVEIEGPONATVYKGSQARFNTVSGMKLIMALSDMWVLSTV 60
 QY 61 RPEPIITNDRTSQRYDQGNFTSEMIHNVPSDSGNIRSLQNSRLHGSAYLTVOYM 120
 DB 61 RPEPIITNDRTSQRYDQGNFTSEMIHNVPSDSGNIRSLQNSRLHGSAYLTVOYM 120
 QY 121 GELFISVNVVAVNEPCEVTCIPSHWTRLPDISWELGLVSHSSYFVPEPSDQSAVS 180
 DB 121 GELFISVNVVAVNEPCEVTCIPSHWTRLPDISWELGLVSHSSYFVPEPSDQSAVS 180
 QY 181 ILALFQSNQTLTCVATWKSILKRSATVNLATYIRCPDPTGGGINIPVLSLPSIGFSL 240
 DB 181 ILALFQSNQTLTCVATWKSILKRSATVNLATYIRCPDPTGGGINIPVLSLPSIGFSL 240
 QY 241 PTMGKVGGLAGTMTLPTCTLTIRCCCRRCRCGCCGCCRCFCRRRRGRIFQKKS 300
 DB 241 PTMGKVGGLAGTMTLPTCTLTIRCCCRRCRCGCCGCCRCFCRRRRGRIFQKKS 300

QY 301 EKEKTKETETESGNGNSGNSDQKTTETASLPKSCSSDPEQNSCGPFPADQR 360
 DB 301 EKEKTKETETESGNGNSGNSDQKTTETASLPKSCSSDPEQNSCGPFPADQR 360
 QY 361 PPRPASHPOASFNINASPEKVSNTTVV 386
 DB 361 PPRPASHPOASFNINASPEKVSNTTVV 386

RESULT 2

AAU75540
 ID AAU75540 standard; Protein; 382 AA.
 AC AAU75540;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Human B7-Like protein, B7-L_{h1}.
 XX
 KW Human: B7-Like protein; B7-L; antiinfectility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antiinfective;
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; anticancer; antifungal;
 KW antiautomatic; nephroprotective; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PE 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX
 DR WPI: 2002-130881/17.
 DR N-PSDB: ABK13030.
 XX
 PT New B7-Like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 13; Fig 1; 135pp; English.

The invention relates to an isolated B7-Like (B7-L) polypeptide (1).
 CC The polypeptide, polynucleotide encoding it and antibody against (1) are
 CC useful for treating B7-Like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (1) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (1) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, Rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as

DB 133 ANEPCFETCLPSHWTWLPDISMELGLVSHSSYFVEPESDLSQAVSILALTPQSNGL 192
 QY 193 TCVATWKSILKARKSATVNLVTRPCPDGTGGINIPGVLSLPSLGFSLPTWKGVLGLAG 252
 DB 193 TCVATWKSILKARKSATVNLVTRPCPDGTGGINIPGVLSLPSLGFSLPTWKGVLGLAG 252
 QY 253 TMLLPTCTLTIRCCCRRCGCCGCCRCRRKGRFRIOFKSEKTKETETE 312
 DB 253 TMLLPTCTLTIRCCCRRCGCCGCCRCRRKGRFRIOFKSEKTKETETE 312
 QY 313 SGNENSGYNSDEOKTETASLPKSCSSDEPORNSSCGPPHORADQPPRASHPOASF 372
 DB 313 SGNENSGYNSDEOKTETASLPKSCSSDEPORNSSCGPPHORADQPPRASHPOASF 372
 QY 373 NLASPEKVSNTTVV 386
 DB 373 NLASPEKVSNTTVV 386

RESULT 4

AAU75543

ID AAU75543 standard; Protein; 377 AA.

AC AAU75543;

DT 23-APR-2002 (first entry)

DE Human B7-like protein, B7-L_{h4}.

Human; B7-like protein; B7-L; anti-infectivity; gynaecological;
 antitumour; cytostatic; immunosuppressive; antiallergic; antineumatic;
 anti-inflammatory; dermatological; antiparasitic; neuroprotective;
 antidiabetic; haemostatic; antithyroid; antituber; antiallergic;
 antiallergic; nephrotropic; antibacterial; vitruoid; tumour; cancer;
 reproductive disorder; graft versus host disease; autoimmune disease;
 toxic shock syndrome; allergy; nephropathy; skin disorder;
 endocrinopathy; lymphoproliferative disorder.

KW Homo sapiens.

PN WO200200710-A2.

PD 03-JAN-2002.

PF 28-JUN-2001; 2001WO-US20719.

PR 28-JUN-2000; 2000US-214512P.

PR 28-NOV-2000; 2000US-0729264.

PA (AMGE-) AMGEN INC.

PI Welher AA, Sarmiento UM, Schultz HJ, Chute HT;

PI WPI: 2002-130881/17.

DR N-PSDB; ABK131.

PT New B7-like polypeptides, polynucleotides and their modulators, useful
 for diagnosing, preventing and treating reproductive, immune and
 proliferative disorders, e.g. cancer and arteriosclerosis -

PS Claim 13; Fig 4; 135pp; English.

CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease disorders or
 CC conditions including reproductive disorders (e.g. infertility, and
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC extracellular disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L

CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or all sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of human B7-L_{h4}.

SQ Sequence 377 AA;

Query Match 72.98; Score 1514; DB 23; Length 377;

Best Local Similarity 86.28; Pred. No. 1.3e-117; Mismatches 284; Conservative 12; Mismatches 20; Indels 6; Gaps 2;

QY 13 GSGGNEVIEBPQNAVYKSGQAFNCTVSGKLMALSDWVLSVRPEEITINDRE 72
 DB 13 GSGGNEVIEBPQNAVYKSGQAFNCTVSGKLMALSDWVLSVRPEEITINDRE 72
 QY 73 TSQRYDGGNFTSMITHNVEPSDSGNIRCSLNSRLHGSAYLTVQMGELFIPSNLVY 132
 DB 73 TSQRYDGGNFTSMITHNVEPSDSGNIRCSLNSRLHGSAYLTVQMGELFIPSNLVY 132
 QY 133 AENPECEVTCPLPSHWTWLPDISMELGLVSHSSYFVEPESDLSQAVSILALTPQSNGL 192
 DB 133 AENPECEVTCPLPSHWTWLPDISMELGLVSHSSYFVEPESDLSQAVSILALTPQSNGL 192
 QY 193 TCVATWKSILKARKSATVNLVTRPCPDGTGGINIPGVLSLPSLGFSLPTWKGVLGLAG 252
 DB 193 TCVATWKSILKARKSATVNLVTRPCPDGTGGINIPGVLSLPSLGFSLPTWKGVLGLAG 252
 QY 253 TMLLPTCTLTIRCCCRRCGCCGCCRCRRKGRFRIOFKSEKTKETETE 312
 DB 253 TMLLPTCTLTIRCCCRRCGCCGCCRCRRKGRFRIOFKSEKTKETETE 312
 QY 313 SGNENSGYNSDEOKTETASLP 334
 DB 310 T---PATIQMKRPQTPLSLP 328

RESULT 5

ABG28169

AC ABG28169;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #28160.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 Homo sapiens.

PN WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Dmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS92356.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID NO 58528; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AbG00010-ABG3037 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 CC
 XX
 SQ Sequence 463 AA;
 Query Match 63.8%; Score 1326; DB 22; Length 463;
 Best Local Similarity 100.0%; Pred. No. 7.3e-102;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 49 MMALSDMVLVSVRPEPLTNDRTSGRYDGGNFTSEMIHANEPSDSNINRCSIONSR 108
 Db 1 MMALSDMVLVSVRPEPLTNDRTSGRYDGGNFTSEMIHANEPSDSNINRCSIONSR 60
 QY 109 LHGSAYLTVOYMGELTIPSVNLVVAENPECVTLPSHMTLPDISMELGLVSHSSYF 168
 Db 61 LHGSAYLTVOYMGELTIPSVNLVVAENPECVTLPSHMTLPDISMELGLVSHSSYF 120
 QY 169 VPESDLOASVSIATLPQSNGLTCAVTKWSLKARSAVNLTVTRCPDGTGGGINIPG 228
 Db 121 VPESDLOASVSIATLPQSNGLTCAVTKWSLKARSAVNLTVTRCPDGTGGGINIPG 180
 QY 229 VLSSLPISLGFSLPTMGVVGGLAGTMTLLPTCTLTIRCCCRRCGCCNCCRCFCRR 288
 Db 181 VLSSLPISLGFSLPTMGVVGGLAGTMTLLPTCTLTIRCCCRRCGCCNCCRCFCRR 240
 QY 289 KRG 291
 Db 241 KRG 243
 RESULT 6

AAU75544
 ID AAU75544 standard; Protein; 370 AA.
 XX
 AC AAU75544;
 XX
 DT 23-APR-2002 (first entry)
 DE Mouse B7-like protein, B7-Lm1.
 XX
 KW Mouse; B7-like protein; B7-L; antiinfectility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antitubercular; antineoplastic;
 KW antiinflammatory; dermatological; antiparasitic; neuroprotective;
 KW antidiabetic; haemostatic; antihypertensive; antileukemic; antiallergic;
 KW antisthmatic; nephrotropic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 XX
 OS Mus musculus.
 PN WO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX
 DR WPI: 2002-130981/17.
 DR N-PSDB; ABX13032.
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 PS Claim 13; Fig 5; 135pp; English.
 XX
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding (I) and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and

CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m1}.

XX Sequence 370 AA:

Query Match 44.7%; Score 928.5; DB 23; Length 370;
 Best Local Similarity 50.0%; Pred. No. 6.4e-69;
 Matches 194; Conservative 54; Mismatches 105; Indels 35; Gaps 5;

QY 5 LLTPPAVSGSGNEVEEQNAVTKGQARNCYVSGKMLIMALSDMYLVSRPME 64
 DB 12 LVLLAQLTASGSSYQIIEBPQNTVLKDEAHNCVTGTGKMLMTLQNMVLSLTQOG 71
 QY 65 PITNDRFTSQRYDQGNFTSEMIHNPSPDSGNIRCSLONSRLHGSAYLVYVWGELE 124
 DB 72 PITNDRFTSQRYDQGNFTSEMIHNPSPDSGNIRCSLONSRLHGSAYLVYVWGELE 131
 QY 125 IPSVNLVAENEPCEVTCPSHWTLPDISWEIGLLVSHSYFYFPEPSDLOSAVSIATL 184
 DB 132 IPSVNLVAENEPCEVTCPSHWTLPDISWEIGLLVSHSYFYFPEPSDLOSAVSIATL 191
 QY 185 TPQSNGLTCVATKMKLARKKATVNLVYIRCPDQGGGINIPVLSLSPGSPPTMG 244
 DB 192 TPQSNGLTCVATKMKLARKKATVNLVYIRCPDQGGGINIPVLSLSPGSPPTMG 239
 QY 245 KVGIGLAGMMLLPPTCLIRCCCRRCGCCGCCRCRRKGRFIDQ---KKSE 301
 DB 240 ILLAAVAFBLILLVLIIFCC-----CASRRKESTYQNEIRKSA 284
 QY 302 KEKTNK---ETETSGENSGINSDEQKTTETASLPKSCSSDPERNSGCPPHRAD 358
 DB 285 NMRTNKADETKLKSKEVGYSSDEAKAQAATASLPKSAEVLSEKRSSTL--PYOELN 342
 QY 359 QRPRASHPOASFNLASPEXSNNTTVV 386
 DB 343 KHOPGPAHPRVSEDIASPOKRVNNTLV 370

RESULT 7
 ID AAU75547 standard; Protein; 631 AA.

XX AAU75547;

DT 23-APR-2002 (first entry)

XX Rat B7-1-like protein, B7-1.

XX Rat; B7-1-like protein; B7-L; antiinferility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antitarratic; antineumatic;
 KW antiinflammatory; dermatological; antiporiatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; antituber; antiallergic;
 KW antiautomatic; nephrotoxic; antibacterial; vitruode; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.

XX Ratus rattus.

PN WO200200710-A2.

XX 03-JAN-2002.

PF 28-JUN-2001; 2001WO-US20719.

XX 28-JUN-2000; 2000US-214512P.

PR 28-NOV-2000; 2000US-0729264.

XX (AMGE-) AMGEN INC.

PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX WPI; 2002-130881/17.
 XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis

XX Disclosure; Fig 8; 135pp; English.

CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allsensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of rat B7-1.

XX Sequence 631 AA:

Query Match 28.3%; Score 587.5; DB 23; Length 631;
 Best Local Similarity 26.4%; Pred. No. 3e-40;
 Matches 170; Conservative 47; Mismatches 112; Indels 315; Gaps 10;

QY 5 LLTPPAVSGSGNEVEEQNAVTKGQARNCYVSGKMLIMALSDMYLVSRPME 25
 DB 12 LVLLAQLTASGSSYQIIEBPQNTVLKDEAHNCVTGTGKMLMTLQNMVLSLTQOG 71
 QY 26 -----NATVKGSGQARFNTVSGKMLIMALSDMYLVSRPMEPITNDRF----- 72
 DB 72 VGVYKSKVRNTVTLKDEAHNCVTGTGKMLMTLQNMVLSLTQOGPITNDRF----- 131
 QY 73 -----TSQRYDQGNFTSEMI 88
 DB 132 LSCDYKFCSEFOSIHRITQMKHDKMVLVSIGVEVMPKRYKNTVYASYNSTDSFISELI 191
 QY 89 IHANPEPSDSGNIRCSLONSRLHGSAYL----- 115
 DB 192 IHDVQPSDSGSSVQCSLHSGFSAFLSVQYVDIANNYSFLLGLILSDRGTYTCVQVR 251
 QY 116 -----TVQVWGELEFIPSVNLVAENEPCEVTCPSHWTLPDISWEIGLLVSHS 164
 DB 252 EGGSYVVKHLITVYVWGELEFIPSVNLVTEGECNVTCAVWGTSPLDISWEILEVPSHS 311

QY 165 -----SYFVPEPSDQ 176
 Db 312 LSVADAPPTNTEYGNPSADIKRTCFASGCPKPLMLNGRELNSYNFLEPGNM 371
 QY 177 SAVSILALTPQSNGLTLCVATWKSARKSATVNL-----211
 Db 372 RVLSTVLDLTPGLNGTLCVAVELKDLQASKSLTVNLGINTTISDPESELYTSSOLDPNA 431
 QY 212 -----TVIRCPDQDGGGINIPGVSLSPLSGFSLPTMGKQVGL 250
 Db 432 TYDHFIDCFIEYGAHYSQNFYVQPPD-----STGECPALPTVAIILLAY 479
 QY 251 AGTML-----TPR 259
 Db 480 AFSILLITVLIITFWKPEDEPDEKQTVPAWAGPAVAKAIIIFAITVIVAVIA 539
 QY 260 CTLVIRCCRRRCGCCGCCRCFCRRRGRIFQO--KSEKEKTKN---ETESG 315
 Db 540 IATIIFFCCCSR-----REKEESTYONEIKRSANMRTNKADPETKLSGK 585
 QY 316 ENSGYSNDEKTEETASLPKSCSESDPEQNSCGPFPORADO 359
 Db 586 ENYSSDDECTVFRRCFRRRNASRETNNKILIGVEAADAQ 629
 RESULT 8
 AAU75545
 ID AAU75545 standard; Protein: 270 AA.
 AC AAU75545;
 DT 23-APR-2002 (first entry)
 DE Mouse B7-1-like protein, B7-L_{m2}.
 XX
 XX Mouse; B7-1-like protein; B7-L; antiinfectivity; gynaecological;
 KM antitumor; cytostatic; immunosuppressive; antirheumatic;
 KM antiinflammatory; dermatologic; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antidiabetic; antiallergic;
 KM antistimulant; nephrotoxic; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; nephropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder.
 XX
 OS Mus musculus.
 XX
 PN MO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001MO-0520719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 PI Welcher AA, Sarmiento UM, Schultz HD, Chute HT;
 XX WPI: 2002-130881/17.
 DR N-PSDB; ABR13033.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 13: Fig 6: 135pp: English.
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and

CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allodermatitis due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m2}.
 XX
 SQ Sequence 270 AA;

Query Match 27.9%; Score 579.5; DB 23; Length 270;
 Best Local Similarity 34.3%; Pred. No. 4.7e-40;
 Matches 133; Conservative 46; Mismatches 74; Indels 135; Gaps 5;

QY 5 LITVPAVSGSGNEVEIEGPNATVILKSGQARFNTVSGKMLMADSVLTVSRPME 64
 Db 12 LVTLAQLTSSGSSYQIIEBPQVNTYVKKSEAHFNCTVTHGKLLMTLNQVYLSLTQGG 71
 QY 65 PTTNDRTSQRDOGNFTSEMIHNPSSDGNIRCSLQNSRLHGSAYLVYVYGELEF 124
 Db 72 PTTNRRFTYASYNSTDSIFSELIIHDVQPSDGSVQCSLQNSHGFGSAFLSVQ----- 125
 QY 125 IPSVNLVAENPECVTCLPSHTRLPDISWELGLVSHSSYTFPEPSDQASVITLAL 184
 Db 126 ----- 125
 QY 185 TPQSNGLTLCVATWKSARKSATVNLTVIRCPDQDGGGINIPGVSLSPLSGFSLPTWG 244
 Db 126 -----DITGEGPALPTWA 139
 QY 245 KVLGLAGTMLTPCTLTIRCCCRRCGCCGCCRCFCRRRGRIFQO---KSE 301
 Db 140 IILLAVAFSLILLIIVLIIFCC-----CASRREKEESTYONEIRKSA 184
 QY 302 KEKTKN---ETESGNSGNSGNSDQKTEETASLPKSCSESDPEQNSCGPFPORADO 358
 Db 185 NMRTNKADPETKLSGKSGEYSSDEAKNAQVSLPSPSAEVSILPEKSSSL--PYQELN 242
 QY 359 GRPPRASHPOASFMILASPEKVSNTTV 386
 Db 243 KHPGPAHFRVSPDIASQKRVNTLV 270

RESULT 9
 AAU75546
 ID AAU75546 standard; Protein: 223 AA.
 AC AAU75546;
 XX

DT 23-APR-2002 (first entry)
 XX
 DE Mouse B7-1-like protein, B7-L_{m3}.
 XX
 KW Mouse; B7-1-like protein; B7-L; antifertility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antileukemic;
 KW antinflammatory; dermatological; antiproliferative; neuroprotective;
 KW antidiabetic; haemostatic; antihypertensive; antitumor; antiallergic;
 KW antiaschemic; nephroprotective; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder.
 XX
 OS Mus musculus.
 XX
 PN MO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001MO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Samlento UM, Schultz HJ, Chute HT;
 XX WPI: 2002-130881/17.
 DR N-PSDB: ABX13034.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis -
 XX
 PS Claim 13; Fig 7; 135pp; English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (1).
 CC The polypeptide, polynucleotide encoding it and antibody against (1) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility, and
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (1) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (1) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or hypersensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the amino acid sequence of mouse B7-L_{m3}.

XX
 SQ Sequence 223 AA;
 Query Match 24.0%; Score 498; DB 23; Length 223;
 Best Local Similarity 31.2%; Pred. No. 2.2e-33;
 Matches 120; Conservative 41; Mismatches 48; Indels 176; Gaps 6;
 QY 5 LITVPEAVSGSGNEVIEGPQNAITVLKSGQARENCYVSGKLTMMALSDMVLVSRPME 64
 DB 12 LVIAQLFLASSSSVQIEGPQNTVLKDSFAHNCVTHGKMLMTLMQMVLLVLTGG 71
 QY 65 PIITNDFRTSORDGCGFTSEMITHVPEPSDSCGNTRCSLQNSRLGSAITVQVNGELF 124
 DB 72 PIITNDFRTVASYNSTDSFISELTHDVQPSDGSVQCSLQNSHGFSAFLSVQ----- 125
 QY 125 IPSVNLVAENEPCEVYCLPSHMTRLPDIDSMELGLVSHSYFVPEPDSQSAVSLAL 184
 DB 126 -----ESTY-----ONEI----- 133
 QY 185 TPQSGTILTCVATWKSILKRSATVNLTVRCPDGTGGINIPGLVSLPSLGSPLTWG 244
 DB 134 -----RKSANM----- 139
 QY 245 KVLGLAGTMLLPCTGLTTRCCCRRCGCCMCRCRCCRRKGRFRIQFKSKSEK 304
 DB 140 -----R 140
 QY 305 TNK---ETETESGNGENSGVNSDEQKTEFTASLPKSCSSDPQRNSSCGPPHORADRP 361
 DB 141 TNKADPETKLKSGKENYSGSDKAQAOTASLPKSAEVSLEPKRSSL--PIQELNKHQ 198
 QY 362 PRPASHPOASENLASPEKVSNTTVV 386
 DB 199 PGATPHRVSFSDIASPQKVRNVTLV 223
 RESULT 10
 AAB81925
 ID AAB81925 standard; protein; 404 AA.
 XX
 AC AAB81925;
 XX
 DT 15-JUN-2001 (first entry)
 XX
 DE Extracorporeal circulation material receptor protein.
 XX
 KW Extracorporeal circulation; carbonyl stress product; receptor;
 KW diabetes; vascular lesion; excretory dysfunction.
 KW
 OS Unidentified.
 XX
 PN WO200118060-A1.
 XX
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000WO-JP06172.
 XX
 PR 08-SEP-1999; 99JP-0254463.
 XX
 PA (TORA) TORAY IND INC.
 XX
 PI Shimizu S, Kubota M, Akiyama H, Usui M;
 XX WPI: 2001-290314/30.
 DR
 XX
 PT Material for extracorporeal circulation, applicable in selective
 PT elimination of diabetic complication factors such as carbonyl stress
 PT products caused by abnormally promoted carbonyl stress from excretory
 PT dysfunction in vascular lesions -
 XX
 PS Claim 1; Page 31-32; 36pp; Japanese.
 XX
 CC The present invention describes a material for extracorporeal circulation

CC which is made from a water-insoluble carrier immobilized with a protein
 CC having the sequence shown here. The materials of the invention, including
 CC adsorbents, are for extracorporeal circulation, which are applicable in
 CC the selective elimination of diabetic complication factors from a body
 CC fluid, and are therefore useful in treating vascular lesions like
 CC arteriosclerosis due to carbonyl stress products caused by abnormally
 CC promoted carbonyl stress from excretory dysfunction.

XX Sequence 404 AA:

Query Match 7.3% Score 151; DB 22; Length 404;
 Best Local Similarity 23.3%; Pred. No. 0.00035;

Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19;

QY 9 PEAVSGSGNEVIEGPONATVILKSGQARFNCVTSQG--WKLIMMALSDMVVLSVRMPMP 65
 DB 124 PEIVDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTISWHLDG-----KP 163
 QY 66 IITNDRFTS-----QRYDGGNFT--SEMITIHVPSDSGNIR-----CSLQNSRLHGSAY 114
 DB 164 LVPNEKGVSVKQOTRRHPETGLTLOSELN--VTPARGDPRPTSCSPGLPRHRAL 220
 QY 115 LTVQVNGELFIP--SVNLVVAENEP-----CEVTCLPISHWTLPDISWE 156
 DB 221 RTAPIQPRVMEPVPLEVQLVY--EPEGGAAPGCTVITLCEVPAOPS-----PQIHWM 272
 QY 157 LGLVSHSSYFYPEPSDLSQSAVSIALLTPQSNGLTCVAIWKSLKARKSATVNLTVIRC 216
 DB 273 KD-----GVPLPLPSPVLIILPEICPODQGTSCVATFHSHPGSRANVSISIE- 322
 QY 217 PDGTGGGINIPGVLSLPLGFSLPTMGVVGGLAGTMLT-----PTCLITTRCCCC 269
 DB 323 PGEEG-----PTAGSVGGSGGLGTALALGILGLGTALLIGVILM 363
 QY 270 RRRCCGCCNCCRCFCRCRRKRGFRIOFKKSEKKT--NKETETESGNSGNSDEQKT 327
 DB 364 QRR-----QRRGERKAPENOEERERALEN----- 389
 QY 328 TETASLPKSCSSDPEQRNSCGPP 353
 DB 390 -----QSEPEPAGESSTGCP 404

RESULT 11

ID AAE3219 standard; Protein: 404 AA.

AC AAE3219;

DT 27-AUG-2002 (first entry)

DE Human receptor for advanced glycosylation end product (RAGE) protein.

KW Human; Receptor for advanced glycosylation end product; RAGE; cardiact;
 KW tissue growth; neointimal formation; blood vessel; restenosis; diabetes;
 KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;
 KW transgenic animal; acute thrombotic stroke; venous thrombosis.

OS Homo sapiens.

PN WO200230889-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US32036.

PR 13-OCT-2000; 2000US-0687528.

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;

XX WPI; 2002-426260/45.

DR N-PSDB; AAD36952.

XX Inhibiting new tissue growth or neointimal formation in blood vessels
 PT of subject suffering from diabetes, stroke and preventing restenosis,
 PT comprising administering inhibitor of receptor for advance glycation end
 PT product

PS Disclosure; Page 16; 43pp; English.

CC The invention relates to a method for inhibiting new tissue growth or
 CC neointimal formation in blood vessels in a subject that has experienced
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic
 CC subject. The method comprises administering an inhibitor of receptor for
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new
 CC tissue growth or neointimal formation in subject's blood vessels and
 CC preventing restenosis in the subject. The method is useful for inhibiting
 CC new tissue growth or neointimal formation in blood vessels in a subject
 CC like non-human animal, a transgenic non-human animal or a human suffering
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable
 CC angina, myocardial infarction, abrupt closure following angioplasty or
 CC stent placement, or thrombosis as a result of peripheral vascular surgery
 CC whether a compound inhibits for preventing restenosis and for determining
 CC subject. The present sequence is human receptor for advanced
 CC glycosylation end product (RAGE) protein.

XX Sequence 404 AA:

Query Match 7.3% Score 151; DB 23; Length 404;
 Best Local Similarity 23.3%; Pred. No. 0.00035;

Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19;

QY 9 PEAVSGSGNEVIEGPONATVILKSGQARFNCVTSQG--WKLIMMALSDMVVLSVRMPMP 65
 DB 124 PEIVDSAS--ELTAGVFN-----KVGTCVSEGSYPAGTISWHLDG-----KP 163
 QY 66 IITNDRFTS-----QRYDGGNFT--SEMITIHVPSDSGNIR-----CSLQNSRLHGSAY 114
 DB 164 LVPNEKGVSVKQOTRRHPETGLTLOSELN--VTPARGDPRPTSCSPGLPRHRAL 220
 QY 115 LTVQVNGELFIP--SVNLVVAENEP-----CEVTCLPISHWTLPDISWE 156
 DB 221 RTAPIQPRVMEPVPLEVQLVY--EPEGGAAPGCTVITLCEVPAOPS-----PQIHWM 272
 QY 157 LGLVSHSSYFYPEPSDLSQSAVSIALLTPQSNGLTCVAIWKSLKARKSATVNLTVIRC 216
 DB 273 KD-----GVPLPLPSPVLIILPEICPODQGTSCVATFHSHPGSRANVSISIE- 322
 QY 217 PDGTGGGINIPGVLSLPLGFSLPTMGVVGGLAGTMLT-----PTCLITTRCCCC 269
 DB 323 PGEEG-----PTAGSVGGSGGLGTALALGILGLGTALLIGVILM 363
 QY 270 RRRCCGCCNCCRCFCRCRRKRGFRIOFKKSEKKT--NKETETESGNSGNSDEQKT 327
 DB 364 QRR-----QRRGERKAPENOEERERALEN----- 389
 QY 328 TETASLPKSCSSDPEQRNSCGPP 353
 DB 390 -----QSEPEPAGESSTGCP 404

RESULT 12

ID AAU77543 standard; Protein: 404 AA.

AC AAU77543;

DT 05-JUN-2002 (first entry)

DE Human receptor for advanced glycosylation end product (RAGE).

KW Receptor for advanced glycation end product; RAGE; receptor;
 KW amyloid beta peptide; blood brain barrier; neurovascular stress;

KM	cerebral vasodilation suppressor; cerebral blood flow enhancer;
KL	cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;
KM	Alzheimer's disease; Down's syndrome; head trauma; stroke; human.
XX	Homo sapiens +
OS	WO200214519-A1.
PN	21-FEB-2002.
PD	14-AUG-2001; 2001WO-US25416.
PF	14-AUG-2000; 2000US-0638648.
PR	(UYCO) UNIV COLUMBIA NEW YORK.
PA	Stern DM, Schmidt AM, Yan SD, Zlokovic B;
P1	WPI; 2002-257610/30.
DR	N-PSDB; ABK10856.
XX	Ameliorating neurovascular stress and decreasing cerebral
PT	vasoconstriction in subject suffering from chronic/acute cerebral
PT	amyloid angiopathy, by administering inhibitor of receptor for advanced
PT	glycation endproduct -
XX	Disclosure; Page 16; 68pp; English.
PS	The invention describes a method of ameliorating neurovascular stress,
CC	and decreasing cerebral vasoconstriction in subject suffering from
CC	chronic or acute cerebral amyloid angiopathy, comprising administering
CC	an inhibitor (I) of receptor for advanced glycation end product (RAGE).
CC	(I) inhibits transcytosis of amyloid beta peptides across blood-brain
CC	barrier, thus decreasing cerebral vasoconstriction and increasing
CC	cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
CC	subject, decreasing cerebral vasoconstriction in a transgenic non-human
CC	animal (preferably, transgenic mouse overexpressing mutant human amyloid
CC	beta precursor protein) or a human, suffering from chronic or acute
CC	cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
CC	ameliorating neurovascular stress comprising cerebral amyloid angiopathy
CC	in a subject, where the neurovascular stress is caused by Alzheimer's
CC	disease, aging, Down's syndrome, head trauma or stroke. This is the
CC	amino acid sequence of human receptor for advanced glycation end
CC	product (RAGE) described in the invention.
CC	
XX	Sequence 404 AA:
SQ	
Query Match	7.3%; Score 151; DB 23; Length 404;
Blast Local Similarity	23.3%; Pred. No. 0.00035;
Matches 90; Conservative	38; Mismatches 112; Indels 146; Gaps 19
QY	9 PEAVSGSGSENYIEGPONATVLKGSQARFNCVTYSOG--WKLMALMSDMVLSVRMEP 65 .
Dbl	124 PEIVDSAS--ELTAGVPN-----KYGTCSSESYPAGTSLWHLDG-----RP 163
QY	66 IITNRFTS-----QRYDGGNT--SEMIIHVEPSDSGNIR-----CSLRNLHGSAV 114
Dbl	164 LPVNEGVSVKEDOTRRHPETGLFTLOSELN--VTPARGDPPTTSCSPCLPPRRRL 220
QY	115 LTVQVWGGEFLIP---SVMLVVVAENEP-----CEVTCLPESHTRLPDISWE 156
Dbl	221 RTRAPIDPRWRPEPLDEYOLV---BPBGAVAPGGVTILTCVPAQPS-----POLHMM 272
QY	157 LGLLVSHSYEFPEDSDQSASVIALIPQSNGLTCAVATWSLKARKSATVNLVIIRC 216
Dbl	273 KD-----GVPLPLPSPVLIILEIGDDPGDTGYCATVSSHGPSRAVISLIE- 322
QY	217 PQDTGGGINPVGVLSSPLSGFSLPWMGXGLLAGMLTF-----PCTLTIRCCC 269
Dbl	323 PGEEG-----PTAGSVGGSGLGTLALALILGLGLTAALLITGIYM 363
QY	270 RRBCCCNCCRCFCRCRRRGFRIOFKSKSEKT--NKETTESNGNSGYNSDEQKT 327

```

Db 364 QRR-----QRGEERKAPNQEERERELN----- 389
QY 328 TETASLPKSCSSDPEGRNSSCCPP 353
Db 390 -----QSEPEAGESSTGGP 404

RESULT 13
AAM48745
ID AAM48745 standard; protein; 404 AA.
XX
XX AAM48745;
XX
XX 02-APR-2002 (first entry)
XX
XX Human RAGE protein SEQ ID NO 1.
XX
XX Human; RAGE; receptor for advanced glycosylated endproduct; receptor;
XX antidiabetic; neuroprotective; cytosolic; antiinflammatory; vasotropic;
XX nephrotropic; dermatological; antiatherosclerotic; noctropic; diabetes;
XX Alzheimer's disease; cancer; inflammation; kidney failure;
XX systemic lupus; nephritis; erectile dysfunction; atherosclerosis.
XX
XX Homo sapiens.
XX
XX WO200192892-A2.
XX
XX 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17447.
XX
XX 30-MAY-2000; 2000US-207342P.
XX
XX 05-MAR-2001; 2001US-0799152.
XX
XX (TRAN-) TRANS TECH PHARMA.
XX
XX Shabaz M;
XX
XX WPI; 2002-114372/15.
XX
XX Detecting a receptor for advanced glycosylated endproducts (RAGE)
XX modulators, for treating e.g., cancer, diabetes or inflammation,
XX comprises measuring the amount of bound anti-RAGE antibody -
XX
XX Claim 1; Fig 2; 49pp; English.
XX
XX The invention relates to detecting receptor for advanced glycosylated
XX endproducts (RAGE) modulators comprises determining the amount of RAGE
XX protein or its fragment bound to the pre-adsorbed ligand by measuring the
XX amount of anti-RAGE antibody bound to the solid surface. The method is
XX useful for rapid, high-throughput identification of compounds that
XX modulate RAGE. The compounds are useful for treating symptoms of diabetes
XX and symptoms of diabetic late complications, amyloidosis, Alzheimer's
XX disease, cancer, inflammation, kidney failure, systemic lupus nephritis
XX or inflammatory lupus nephritis, erectile dysfunction and
XX atherosclerosis.
XX
XX Sequence 404 AA:
XX
XX Query Match 7.3%; Score 151; DB 23; Length 404;
XX Best Local Similarity 23.3%; Pred. No. 0.00035;
XX Matches 90; Conservative 38; Mismatches 112; Indels 146; Gaps 19
QY 9 PEAVSGSGNEVIEGQNAITVLKGSARFCVTYSOG---WKLIMALSDMVVLSVRPMP 65
Db 124 PELITVDAS--ELTAGVPN-----KGTCTVSEGSYAGTILSHIDG-----KP 163
QY 66 IITNDRFTS-----QRYDQGNFT--SEMIHNVEPSDSGNIR---CSLQNSRLHGSAV 114
Db 164 LTVNENGVSGKEQTRRHRETGLFTLSLMT--VTBARCGDPRPTSCFSFPLPRHRL 220
QY 115 LTVNENGVGLFIP-----SVNLVYANEP-----CVYGLPBGHWITPLPISWE 156

```

Db 221 RTAPLOPRWEPVPEEVLVV---EPEGAVAPGTVLTCVEPAOS-----PQIHMM 272
 QY 157 LGLVSHSSSYFVPEPSDLSQASVIALTPQSNGLTCVATWKSILKARSAVNLTVIRC 216
 Db 273 KD-----GVPLPLPSPVLLPEIFPQOGITYSCVATHSHGPOSRNAVSTISLE- 322
 QY 217 PODTGGGINIPGVLSLPSLPGSLPTWGRVGLAGTMLLT-----PTCTLTTRCCC 269
 Db 323 PDEEG-----PTAGSVGGSGGLTALAGLIGLGTALLGVILM 363
 QY 270 RRRCCGNCNCCHRCFCRRKRGRFRIOFQKSEKKT--NKETESGNGNSGYNDEOKT 327
 Db 364 QRR-----ORGEERRAPENOEERAEIN----- 389
 QY 328 TETASLPKSCSSDPEORNSCGPP 353
 Db 390 -----QSEPEAGESSTGCP 404
 .RESULT 14
 AAM81030
 ID AAM81030 standard; Protein: 1496 AA.
 AC AAM81030;
 XX
 DT 10-MAY-1999 (first entry)
 XX
 DE Melanoma associated antigen MG50.
 KM MG50; melanoma gene-50; melanoma associated antigen; human;
 KW T cell epitope; cancer; lung cancer; rhabdomyosarcoma; diagnosis;
 KW therapy; vaccine.
 OS Homo sapiens.
 XX
 PN W0985513-A1.
 XX
 PD 10-DEC-1998.
 XX
 PF 04-JUN-1998; 98WO-US11533.
 XX
 PR 06-JUN-1997; 97US-0870941.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (URS-) UNIV SOUTHERN CALIFORNIA.
 XX
 P1 Deans RJ, Kap-Mitchell J, Minev BR, Mitchell MS;
 XX
 DR N-PSDB; AAV99922.
 DR
 XX
 PS Claim 1; Page 45-49; 79pp; English.
 XX
 CC This polypeptide comprises a portion of a new human melanoma
 CC associated antigen, designated MG50. The amino acid sequence was
 CC deduced from a cDNA clone (see AAV99922) isolated from melanoma
 CC cell line MSW M-1 cDNA by subtractive hybridisation. The 5' region
 CC of MG50 cDNA was not obtained. MG50 mRNA has been detected in
 CC melanoma, lung carcinoma and rhabdomyosarcoma cells, foetal brain,
 CC foetal heart and human placenta. The invention also provides T
 CC cell epitopes (see AAM81031-54) from MG50, including cytotoxic and
 CC helper T cell epitopes, antibodies that specifically bind to MG50
 CC or an MG50 T cell epitope, recombinant vectors, and antigen
 CC presenting cells. Methods are provided for identifying an MG50
 CC melanoma associated antigen in an individual and for identifying an
 CC immune response against an MG50 melanoma associated antigen, as
 CC well as methods of stimulating T lymphocytes that are reactive
 CC against cancer cells expressing MG50 and for treating an individual

CC having cancer cells that express MG50. The products and methods
 CC can be used for the detection, treatment and prevention of
 CC MG50-expressing cancers, e.g. melanomas, lung cancer or
 CC rhabdomyosarcoma.
 CC
 SQ Sequence 1496 AA;
 Query Match 6.5%; Score 135; DB 20; Length 1496;
 Best Local Similarity 25.0%; Pred. No. 0.041;
 Matches 73; Conservative 36; Mismatches 127; Indels 56; Gaps 15;
 QY 13 GSGSGNEVIEGPONATVLKGSQARFNCTVS-QGKLIIMALSDMVVSVPMPEITNDR 71
 Db 354 GSPARPTFVIGDPQNTVELVGSVTLKESAGHPPRRISWTRGDRTPLPVPRNNIPS-- 411
 QY 72 FTSRTOGCFNFTSEMIIHNVEPSDSGNIRCSLONS--RLHGSAYLVTVQMGELFIPSVN 129
 Db 412 -----GG-----LYIONVQDGSGEVACSAATNNIDSVHATAFTIVCALPQFTVPOD 458
 QY 130 LVVAENPECEVTCLEPSHMTRLPDISM-ELGLVSHSSSYFVPEPSDLSQASVIALTPQS 188
 Db 459 RVVIEGOTVDFQC-EAKGNPPVPIAMTKGSSQLSVDRRLVLSSGTLR--ISGVALLDQ- 514
 QY 189 NGTLVCVATWKSILKARSAVNLTV-----IRCPQDT-----GGGINIPGVLSLPSL 236
 Db 515 -GQYECQAV--NIIGSKVVAHLTVQPRVPPVFPASIPSDITVEGVANVQLP-----CSSQ 566
 QY 237 GFSLP--TWGRVIGL--AGTMLTPTCTLT-----RCCCRRCRCCG 275
 Db 567 GEPPALTMKDGQVYESGRFHSIPGFLTLINDVGPADAGRECVARNITG 618
 .RESULT 15
 AAY70469
 ID AAY70469 standard; Protein: 1496 AA.
 AC AAY70469;
 XX
 DT 21-JUN-2000 (first entry)
 XX
 DE Human p53 target molecule, PRG2 protein.
 XX
 KW PRG2; p53 target; human; modulate; cell proliferation; immunomodulatory;
 KW chromosome 2p24.3; cytostatic; gene therapy; tumour cell; inducer;
 KW diagnosis; therapeutic; proliferative disease; cell cycle arrest; cancer;
 KW treatment; apoptosis; knockout animal; cancer susceptibility; dxn; hpxn;
 KW peroxidase; redox-regulation; reactive oxygen species; ROS.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..29
 FT /label= Signal-peptide
 FT Protein 30..1496
 FT /note= "Mature_human_PRG2_protein
 FT /note= "Homologue of Drosophila peroxidase, dxn"
 XX
 PN W0200012526-A1.
 PN
 XX
 PD 09-MAR-2000.
 XX
 PF 27-AUG-1999; 99WO-US19551.
 XX
 PR 28-AUG-1998; 98US-0098251.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 P1 Horikoshi N, Shenk T;
 XX
 DR WPI: 2000-246724/21.
 DR N-PSDB; AAK51671.
 XX
 PT New p53-inducible isolated nucleic acid molecule including open reading

PT frame encoding human homolog of Drosophila melanogaster peroxidase,
 PT useful e.g. in detection and treatment of cancer -

PS Disclosure; Page 71-74; 83pp; English.

XX
 CC The present sequence is the protein encoded by PRG2 gene, that is
 CC upregulated in response to induction of p53 activity in human colon
 CC cancer EBI cells. This sequence is the human homologue (hpxn) of
 CC Drosophila peroxidase gene dpxn, that is expressed in heart, placenta,
 CC spleen, ovary and intestines. PRG2 is involved in p53-mediated growth
 CC suppression pathways and plays a role in redox regulation. It is a
 CC haem-peroxidase that increases the intracellular content of reactive
 CC oxygen species (ROS). They are potential targets of p53 regulatory
 CC activity and are useful for modulation of cellular proliferation. PRG2
 CC gene is localised to human chromosome 2p24.3. The PRG target molecules
 CC have cytostatic and immunomodulatory activity. PRG polynucleotides,
 CC proteins and antibodies are useful as diagnostic and therapeutic agents
 CC for detection and treatment of cancer and other proliferative diseases.
 CC The gene/cDNA may be used for gene therapy, to restore a gene function
 CC downstream of p53, that cannot be activated in the p53-deficient tumour
 CC cell. Antibodies can be used as inducers of cell cycle arrest and/or
 CC apoptosis. The DNA sequences can be used to generate 'knockout' animals
 CC as a model of cancer susceptibility.

SQ Sequence 1496 AA:

Query Match 6.5%: Score 135; DB 21; Length 1496;

Best Local Similarity 25.0%: Pred. No. 0.041;
 Matches 73; Conservative 36; Mismatches 127; Indels 56; Gaps 15;

QY 13 GSGSGNEVIEGPNATVILKGSQARFNCTVS-QGKMLIMALSDMVILSVRPMPEITNDR 71
 DB 354 GSPARPTFVIOQNTVEVLVGESVTLCSATGHPPRISMTRGDRPLPVDPRVNITPS-- 411
 QY 72 FTSQRYDOGNTSMITHNPEPSDGNIRCSLONS--RLGSAVLYQVVGELFIPSVN 129
 DB 412 -----GG-----LYIQNVVQSGSGEYACSAIWNNDISVAHTAFIVQALPQFTVTPOD 458
 QY 130 LVVAENPECEVTCPLPSHWTPLPDISM-ELGLVSHSSYYFPEPSDLOSASVILATPQS 188
 DB 459 RVIEGQVVDPCQ-EAKGNPPVIAWTKGSGQLSVDRRLVLSGTLR--ISGVALHDQ- 514
 QY 189 NCTLCVATWKSILKARKSATVNLIV-----IRCPQDT-----GGINIPGVLSLPSL 236
 DB 515 -GOYECCQAV--NIISQKVAHLFVQPRVTPVFASIPSDTVEVGANVQLP-----CSSQ 566
 QY 237 GFSLP--TWGKVGGLT--AGTMLLPPTCTLT-----RCCCCRRCRG 275
 DB 567 GEPEPATWKNKDGVOVTESGKFHISPEGLTINDVGPADAGRICEVARNITG 618

Search completed: April 28, 2003, 18:14:34
 Job time : 41.4662 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 23:25:02 ; Search time 1905.64 Seconds
(without alignments)
9985.963 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 ctgtctgccatctgaataa.....gtaatacaactgtatag 1175

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estbta:*
2: em_estbhm:*
3: em_estlin:*
4: em_estlmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Match Length	DB ID	Description
1	308.6	26.3	784 12	BG863468
2	279	23.7	794 12	BF784177
3	267.4	22.8	725 12	BG740428
4	259.4	22.1	631 10	BB663870
5	247.8	21.1	527 10	BE032610
6	199	16.9	517 10	BE476432

7	191.6	16.3	524 10	AV601192	AV601192
8	190.8	16.2	622 17	AG069679	AG069679
9	190.8	16.2	677 17	AG107877	AG107877
10	186	15.8	471 9	AA265274	AA265274
11	186	15.8	911 10	AW990468	AW990468
12	186	15.8	631 12	BG173684	BG173684
13	186	15.8	1552 11	AK008060	AK008060
14	185	15.7	466 9	AI425363	AI425363
15	184.4	15.7	663 10	BE376580	BE376580
16	182.8	15.6	650 10	BE625217	BE625217
17	168.2	14.3	600 17	AZ379623	AZ379623
18	153.2	13.0	216 12	BG206666	BG206666
19	153.2	13.0	556 9	AI788300	AI788300
20	151.6	12.9	530 9	AI790785	AI790785
21	145.8	12.4	423 10	BB846577	BB846577
22	143.8	12.2	291 17	AZ411779	AZ411779
23	128.8	11.0	471 12	BF040046	BF040046
24	127.4	10.8	394 10	BB846133	BB846133
25	124	10.6	685 17	AG142221	AG142221
26	110	9.4	785 13	BI452873	BI452873
27	102.8	8.7	889 13	BI454276	BI454276
28	100	8.5	185 14	N47851	N47851
29	100	8.5	234 14	N93995	N93995
30	98.4	8.4	525 17	AQ403719	AQ403719
31	78.6	6.7	360 9	A1647601	A1647601
32	78.6	6.7	374 9	A1648720	A1648720
33	78.6	6.7	406 9	A1788973	A1788973
34	77.7	6.6	644 10	BB531388	BB531388
35	67.8	5.8	397 10	AW112084	AW112084
36	63.6	5.4	282 10	BB564363	BB564363
37	56	4.8	818 12	BG445021	BG445021
38	52.6	4.5	422 9	A1756062	A1756062
39	51	4.3	457 9	AJ003343	AJ003343
40	50.8	4.3	458 9	AL514929	AL514929
41	50	4.3	447 10	AV431233	AV431233
42	49.8	4.2	428 9	AA635064	AA635064
43	49.6	4.2	925 17	CNS0381E	CNS0381E
44	49.6	4.2	942 17	CNS00601	CNS00601
45	49.4	4.2	481 10	BE017134	BE017134

ALIGNMENTS

RESULT 1
BG863468
LOCUS
DEFINITION
602796941F1 NCL_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4917803 5',
mRNA sequence.
ACCESSION
BG863468
VERSION
BG863468.1 GI:14214006
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
REFERENCE
1 (Bases 1 to 784)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
plate: LLAM10830 row: b column: 12
High quality sequence stop: 768.
Location/Qualifiers

FEATURES

source

1. 784
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4917803"
/clone_lib="NCI CGAP Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lotter Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 188 a 202 c 208 g 186 t

Query Match 26.3%; Score 308.6; DB 12; Length 784;
Best Local Similarity 69.5%; Pred. No. 2.3e-61;
Matches 447; Conservative 0; Mismatches 194; Indels 2; Gaps 2;

24 GAGATGGGCTTGATTTCTCCACGGTCTGGCTGTGTAATGAACTAGAGGC 83
123 GTGCTGGTATCTCGCTGAGCTGACAGCTTCGGATTCAGATCATAGAGGT 182
84 CCCCAGATGCAACACTCTGAAAGGCTCCAGGCTCGCTTGAAGTCCCTCCAG 143
183 CCTCAGATGTAACACTCTGAAAGGCTCAGAGGCTCACTTAAGTCCGTAAGT 242
144 GCGTGAAGCTCATCTGTTGGCTGCAAGTGGTGTCTAAGCTCAGGCCATG 203
243 GCGTGAAGCTCTCATGTTGGCTGCAAGTGGTGTCTAAGTCCGTAAGT 302
204 GAGCCATCATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 263
303 GGACCCATCATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 362
264 ACCTGGAGTGTATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 323
363 ATCTGGAGTGTATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 422
324 CTCGAGAGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 383
423 CTCGAGAGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 482
384 TTATTTCCAGCTTATCTGTTAGTGGTGAATGAACTTTGTAAGTACTTCTCA 443
483 AACATTTCTGAGCAACACTTATGATGAGGTGAACCCCTGTAATGTGACTTGTAT 542
444 CCTCAGAGTGTATCTGTTAGTGGTGAATGAACTTTGTAAGTACTTGTAT 503
543 GCGGTGGGCTGAGCTCACTCCCGGATATTTCTGAGAGCTTGAAGTCCCTAAGCAT 602
504 TCAAGCTATTTTGTTCGAGAGCCAGGACCTTCAAGTGCAGTGAAGTCTTGGCT 563
603 TCGAGTAACTATCTTCTGAGAGCCGGAAGTATGAGGCTTGAAGTCTTGGCTGAGC 662
564 CTGAGCCAGAGAGGAGTGGAGTGGTGGTGAAGTGGTGAAGTGGTGAAGTGGC 623
663 CTGAGCCAGAGAGGAGTGGAGTGGTGGTGAAGTGGTGAAGTGGTGAAGTGGC 721
624 CGAAGCTGCAAGTGAATCTGAGTGGTGAAGTGGTGAAGTGGTGAAGTGGC 666
722 AGCAAGTCTTAAGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGC 763

RESULT 2
BF784177 794 bp mRNA linear EST 12-JAN-2001
LOCUS BF784177
DEFINITION 602108039P1 NCI CGAP Kid14 Mus musculus cDNA clone IMAGE:4236489
5', mRNA sequence.
ACCESSION BF784177

VERSION BF784177.1 GI:12089213
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS I (bases 1 to 794)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9844 row: n column: 10
High quality sequence stop: 694.

FEATURES
Location/Qualifiers
1. 794
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4236489"
/clone_lib="NCI CGAP Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 186 a 214 c 215 g 179 t

Query Match 23.7%; Score 279; DB 12; Length 794;
Best Local Similarity 68.3%; Pred. No. 1.8e-54;
Matches 416; Conservative 0; Mismatches 190; Indels 3; Gaps 2;

24 GAGATGGGCTTGATTTCTCCACGGTCTGGCTGTGTAATGAACTAGAGGC 83
123 GTGCTGGTATCTCGCTGAGCTGACAGCTTCGGATTCAGATCATAGAGGT 182
84 CCCCAGATGCAACACTCTGAAAGGCTCCAGGCTCGCTTGAAGTCCCTCCAG 143
183 CCTCAGATGTAACACTCTGAAAGGCTCAGAGGCTCACTTAAGTCCGTAAGT 242
144 GCGTGAAGCTCATCTGTTGGCTGCAAGTGGTGTCTAAGCTCAGGCCATG 203
243 GCGTGAAGCTCTCATGTTGGCTGCAAGTGGTGTCTAAGTCCGTAAGT 302
204 GAGCCATCATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 263
303 GGACCCATCATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 362
264 ACCTGGAGTGTATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 323
363 ATCTGGAGTGTATCAACATGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 422
324 CTCGAGAGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 383
423 CTCGAGAGAGGCTTCACTCTCAGAGGTACGAGCGGGGAGCTTC 482
384 TTATTTCCAGCTTATCTGTTAGTGGTGAATGAACTTTGTAAGTACTTCTCA 443
483 AACATTTCTGAGCAACACTTATGATGAGGTGAACCCCTGTAATGTGACTTGTAT 542
444 CCTCAGAGTGTATCTGTTAGTGGTGAATGAACTTTGTAAGTACTTGTAT 503
543 GCGGTGGGCTGAGCTCACTCCCGGATATTTCTGAGAGCTTGAAGTCCCTAAGCAT 602
504 TCAAGCTATTTTGTTCGAGAGCCAGGACCTTCAAGTGCAGTGAAGTCTTGGCT 563

Db 603 TCGAGTACATTCCTTCTGAGACC--GGGACCTTATGAGGCTGAGTGTCTGGAA 660
 Oy 564 CTTGCCCCAGAGGCAATGGAGCTTGTGACTTGGCTTACCTTGAAGACCTTGAAGGCC 623
 Db 661 CTCA-CACACTGGGCAAGGAGCTTGTGTGGCAGAGCTGAAGACTTGCAGGC 719
 Oy 624 CGCAAGTCT 632
 Db 720 CAGCAGTCT 728

RESULT 3

BG740428 725 bp mRNA linear EST 15-MAY-2001
 LOCUS BG740428
 DEFINITION 60263817F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4778789 5',
 mRNA sequence.

ACCESSION BG740428
 VERSION BG740428.1 GI:14051081
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 725)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL0635 row: b column: 06
 High quality sequence stop: 725.
 Location/Qualifiers
 1. 725

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4778789"
 /clone_lib="NCI_CGAP_Skn3"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP library."
 BASE COUNT 227 a 163 c 159 g 176 t
 ORIGIN

Query Match 22.8%; Score 267.4; DB 12; Length 725;
 Best Local Similarity 99.6%; Pred. No. 8.9e-52;
 Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 906 AAGAAATCTGAAAAAGAGAAACAAGAACTGAGACAGAAGTGAATGAAGAAAC 965
 Db 377 AGGAATCTGAAAAAGAGAAACAAGAACTGAGACAGAAGTGAATGAAGAAAC 436
 Oy 966 TCCGGCTACATTTCAATGATGACAAAAGACACAGACCCGCTTCTCTCCCAATCC 1025
 Db 437 TCCGGCTACATTTCAATGATGACAAAAGACACAGACCCGCTTCTCTCCCAATCC 496
 Oy 1026 TGTGATCAGTGTCTGACAAAGAAAGAGTACTGTGGCCCTCCACACAGGGCT 1085
 Db 497 TGTGATCAGTGTCTGACAAAGAAAGAGTACTGTGGCCCTCCACACAGGGCT 556
 Oy 1086 GATCAAGTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTTTAATCTGGCAGTCT 1145
 Db 557 GATCAAGTCCACCCAGGCGCAGCAAGTATCCACAGGCTTCTTTAATCTGGCAGTCT 616

Oy 1146 GAGAGGTGATATACACTGATATATA 1174
 Db 617 GAGAGGTGATATACACTGATATATA 645

RESULT 4

LOCUS BB663870
 DEFINITION BB663870 RIKEN full-length enriched, 0 day neonate lung Mus
 musculus cDNA clone E030016M12 5', mRNA sequence.

ACCESSION BB663870
 VERSION BB663870.1 GI:16497624
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 631)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
 Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillar sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa,
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
 Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES

source

location/Qualifiers
 1. 631
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_image="E030016M12"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 lung"
 /tissue_type="lung"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia

RESULT 7	524 bp	mRNA	linear	EST 27-NOV-2001
LOCUS				
AV601192	AV601192	Bos taurus	kidney testis	Bos taurus cDNA clone EIK1002H04
DEFINITION	5', mRNA	sequence,		
ACCESSION	AV601192			
VERSION	AV601192.1	GI:9723510		
KEYWORDS				
SOURCE				
COM.				

RESULT	8
LOCUS	AG069679
DEFINITION	Pan troglodytes DNA, clone: PTB-060F12.F, genomic survey sequence.
ACCESSION	AG069679
VERSION	AG069679.1 GI:16621481
KEYWORDS	GSS.
SOURCE	Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC library clone:PTB-060F12.F.

ORGANISM	Pan troglodytes			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
REFERENCE	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.			
REFERENCE	BAC end sequences of library PTB			
REFERENCE	Unpublished			
REFERENCE	2 (bases 1 to 622)			
REFERENCE	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)			
COMMENT	BAC clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.			
PRIMERS	Sequencing: -21M13			
LIBRARY	Vector : pRS145			
	R.Site 1 : SacI			
	R.Site 2 : SacI.			
FEATURES	Location/Qualifiers			
source	1..622			
	/organism="Pan troglodytes"			
	/db_xref="taxon:9598"			
	/clone="PTB-060F12.F"			
	/sex="male"			
	/cell_type="lymphoblast"			
	/clone_lib="PTB Chimpanzee Male BAC Library"			
BASE COUNT	165 a 153 c 128 g 176 t			
ORIGIN				
Query Match	16.2%: Score 190.8; DB 17; Length 622;			
Best Local Similarity	94.3%: Pred. No. 5.8e-34;			
Matches 198; Conservative	0; Mismatches 12; Indels 0; Gaps 0;			
OY	474 TCCTGGAGCTCGGCTCTCTCTGTCAGCCATTCATTTATTTTGTCCGAGCCAGC 533			
Db	72 TCCTGGAGCTCGGCTCTCTCTGTCAGCCATTCATTTATTTTGTCCGAGCCAGC 131			
OY	534 GACCTTCAAAGTGCAGTGCATCTCTGCTCGAGCCACAGAGAAGGGACTTGACT 593			
Db	132 GACCTTCAAAGTGCAGTGCATCTCTGCTCGAGCCACAGAGAAGGGACTTGACT 191			
OY	594 TCGCTGGCTACCTGGAAGAGCCTGAAGGCCGCGACAGTGTGCACACTGTAATCTCACTGTG 653			
Db	192 TCGCTGGCTACCTGGAAGAGCCTGAAGGCCGCGACAGTGTGCACACTGTAATCTCACTGTG 251			
OY	654 ATTGCGTGTCCCAAGACACTGGAGGTGT 683			
Db	252 ATTGCGCGTCCCAAGGTAAAGTGAAGATGT 281			
RESULT 9				
LOCUS	AG107877 677 bp DNA linear GSS 03-NOV-2001			
DEFINITION	Pan troglodytes DNA, clone: PTB-112N07.F, genomic survey sequence.			
ACCESSION	AG107877			
VERSION	AG107877.1 GI:16728395			
KEYWORDS	GSS.			
SOURCE	Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male			
ORGANISM	BAC library clone:PTB-112N07.F.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
REFERENCE	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.			

TITLE	BAC end sequences of Library PTB
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 677)
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Toto, K. I., Watanabe, H. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (02-APR-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/ Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT	Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS	Sequencing: -21M13
LIBRARY	Vector : pKS145
R.Site 1	: SacI
R.Site 2	: SacI.
Location/Qualifiers	1. 677
/organism	= "Pan troglodytes"
/db_xref	= "taxon:9598"
/clone	= "PTB-112N07.F"
/sex	= "male"
/cell_type	= "lymphoblast"
/clone_lib	= "PTB Chimpanzee Male BAC Library"
BASE COUNT	168 a 159 c 143 g 207 t
ORIGIN	
Query Match	16.2% Score 190.8; DB 17; Length 677:
Best Local Similarity	94.3% Pred. No. 5,9e-34; Mismatches 12; Indels 0; Gaps 0;
Matches 198; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
QY 474	TCCTGGAGCTGGTCTCCTGTCGACCATTCAGCTATTTTGTCCGGAGCCCAAC 533
Db 72	TCCGCGAGACTGGTCTCCTGTCGACCATTCAGCTATTTTGTCCGGAGCCCAAC 131
QY 534	GACCTTCAAGTCGACATGACATCTGCTGTCGACCCACAGAGCAATGGACTTGTACT 593
Db 132	GACCTTCAAGTCGACATGACATCTGCTGTCGACCCACAGAGCAATGGACTTGTACT 191
QY 594	TGCGTGGCTACCTGGAAGAGCCTGGAAGGCCCGCAAGTCTGCACTGTAAATCTCACTGTG 653
Db 192	TGCGTGGCTACCTGGAAGAGCCTGGAAGGCCCGCAAGTCTGCACTGTAAATCTCACTGTG 251
QY 654	ATTGCGTGTCCCAAGACACTGGAGGTGCT 683
Db 252	ATTGCGGTCTCCCAAGTGAAGATGT 281
RESULT 10	
AA265274	AA265274 471 bp RNA linear EST 20-MAR-1997
LOCUS	mx91d09.r1 Soares mouse NML Mus musculus cdna clone IMAGE:693713
DEFINITION	5', mRNA sequence.
ACCESSION	AA265274
VERSION	AA265274.1 GI:1901370
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus. 1 (bases 1 to 471)
AUTHORS	Mair, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE	The WashU-BHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Maria M/Mouse EST project

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; R

CTCGGAGATGATCATCCACATGTGGAGCCAGTGATTCCGGGAACATCAGATGCAGC 323

1 234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300

Db 150 CCTGAGATGTAAACAGTCTTAAGAGACTCAGAGGCTCACTTCAACTGCACCGTGACTCAC 209
 Oy 144 GGCTGGAAGCTCATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 203
 Db 210 GGCTGGAAGCTCTCATATGTTGGGCTCTTAAGCAATGTTGGTGTGCTTAAGGCTCAGGCCATG 269
 Oy 204 GAGCCCATCATCAGCAATGAGCCGCTTCACTCTCAGAGGTACGACCGAGGCGGGAATTC 263
 Db 270 GGACCCATCATCAGCAATGAGCCGCTTCACTCTCAGAGGTACGACCGAGGCGGGAATTC 329
 Oy 264 ACCTGGAAGTGTATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 323
 Db 330 ATCTCGAGTGTATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 389
 Oy 324 CTCGAGAACAGTGCCTCATGATGATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 369
 Db 390 CTCGAGAACAGTGCCTCATGATGATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 435

RESULT 15

BE376590

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

BE376590

VERSION

BE376590.1

KEYWORDS

SOURCE

EST.

ORGANISM

house mouse.

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM8743 row: f column: 06

High quality sequence stop: 498.

Location/Qualifiers

1. 663

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3584645"

/clone_lib="NCI-CCGAP_Mam1"

/tissue_type="tumor, biopsy sample"

/dev_stage="10 months, virgin"

/lab_host="DH10B"

/note="Organ: mammary; Vector: PCMV-SPORT6; Site: 1; Salt; Site: 2; Not: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT

ORIGIN

143 a 193 c 170 g 157 t

Query Match 15.7%; Score 184.4; DB 10; Length 663;

Best Local Similarity 70.8%; Pred. No. 1.8e-32;

Matches 245; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Db 125 CCTGAGATGTAAACAGTCTTAAGAGACTCAGAGGCTCACTTCAACTGCACCGTGACTCAC 184
 Oy 144 GGCTGGAAGCTCATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 203
 Db 185 GGCTGGAAGCTCTCATATGTTGGGCTCTTAAGCAATGTTGGTGTGCTTAAGGCTCAGGCCATG 244
 Oy 204 GAGCCCATCATCAGCAATGAGCCGCTTCACTCTCAGAGGTACGACCGAGGCGGGAATTC 263
 Db 245 GGACCCATCATCAGCAATGAGCCGCTTCACTCTCAGAGGTACGACCGAGGCGGGAATTC 304
 Oy 264 ACCTGGAAGTGTATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 323
 Db 305 ATCTCGAGTGTATCATATGTTGGGCTCTCAGTGCATGATGGTGTGCTTAAGGCTCAGGCCATG 364
 Oy 324 CTCGAGAACAGTGCCTCATGATGATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 369
 Db 365 CTCGAGAACAGTGCCTCATGATGATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 410

Search completed: April 29, 2003, 09:24:07

Job time: 1924.64 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
1	269	22.9	474	9	US-09-918-995-3342		Sequence 3342, App
2	190.4	16.2	401	10	US-09-864-761-16305		Sequence 16305, App
3	142	12.1	398	10	US-09-983-965-4945		Sequence 4945, App
4	46	3.9	3577	9	US-10-008-739A-1		Sequence 1, App
5	44.8	3.8	1310	9	US-09-849-243-13		Sequence 13, App
6	44.8	3.8	2614	9	US-09-822-845-421		Sequence 421, App
7	44.8	3.8	3283	9	US-09-849-243-15		Sequence 15, App
8	44.8	3.8	4286	9	US-09-849-243-14		Sequence 14, App
9	44.8	3.8	6604	10	US-09-880-107-1748		Sequence 1748, App
10	44	3.7	455	9	US-09-728-444-151		Sequence 151, App
11	43.4	3.7	384	9	US-09-738-626-544		Sequence 544, App
12	43.4	3.7	3309400	9	US-09-738-626-1		Sequence 1, App
13	43.2	3.7	293	10	US-09-864-761-18923		Sequence 18923, App
14	43.2	3.7	459	10	US-09-864-761-2182		Sequence 2182, App
15	43.2	3.7	1791	10	US-09-416-384A-6		Sequence 6, App
16	42.6	3.6	2424	9	US-10-029-217A-25		Sequence 25, App
17	42.6	3.6	4959	9	US-10-029-217A-1		Sequence 1, App
18	42.6	3.6	4960	9	US-10-029-217A-30		Sequence 30, App
19	42.2	3.6	5959	10	US-09-954-456-1996		Sequence 1996, App

C	20	41.6	3.5	238	10	US-09-864-761-24079	Sequence 24079, A
C	21	41.6	3.5	554	10	US-09-864-761-7357	Sequence 7357, A
C	22	41.6	3.5	3899	10	US-09-735-3678-5	Sequence 5, Appl1
C	23	41.6	3.5	6504	10	US-09-969-347-171	Sequence 171, Appl
C	24	41.6	3.5	6504	10	US-09-735-3678-9	Sequence 9, Appl1
C	25	41.6	3.5	6878	10	US-09-735-3678-1	Sequence 1, Appl1
C	26	41.4	3.5	2572	10	US-09-925-300-486	Sequence 486, Appl
C	27	41.2	3.5	424	9	US-09-918-995-10212	Sequence 10212, A
C	28	41.2	3.5	725	10	US-09-919-580-878	Sequence 878, Appl
C	29	41	3.5	2517	10	US-09-476-242-16	Sequence 16, Appl
C	30	40.8	3.5	2352	10	US-09-476-242-26	Sequence 26, Appl
C	31	40.8	3.5	2336	9	US-09-822-846-129	Sequence 129, Appl
C	32	40.8	3.5	6177	10	US-09-916-145-1	Sequence 1, Appl1
C	33	40.8	3.5	6322	10	US-09-917-8006-1546	Sequence 1546, Appl
C	34	40.6	3.5	547	9	US-09-918-995-11012	Sequence 11012, A
C	35	40.4	3.4	954	9	US-10-078-090-83	Sequence 83, Appl
C	36	40.4	3.4	2674	9	US-10-003-295-1	Sequence 1, Appl1
C	37	40.2	3.4	216	9	US-09-728-444-25	Sequence 25, Appl
C	38	40.2	3.4	545	10	US-09-919-580-371	Sequence 371, Appl
C	39	40.2	3.4	557	10	US-09-919-580-475	Sequence 475, Appl
C	40	40.2	3.4	572	10	US-09-919-580-570	Sequence 570, Appl
C	41	40.2	3.4	1143	9	US-09-938-842A-1359	Sequence 1399, Ap
C	42	40.2	3.4	2523	10	US-09-476-242-15	Sequence 15, Appl
C	43	40.2	3.4	83450	9	US-09-811-469-3	Sequence 3, Appl1
C	44	40	3.4	270	9	US-10-007-557-8	Sequence 8, Appl1
C	45	40	3.4	488	9	US-09-918-995-19918	Sequence 19918, A

ALIGNMENTS

```

RESULT 1
US-09-918-995-3342
: Sequence 3342, Application US/09918995
: Publication No. US20030073623A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918,995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235,076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3342
: LENGTH: 474
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(474)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-3342

```

Query Match	22.98;	Score 269;	DB 9;	Length 474;
Best Local Similarity	98.2%;	Pred. No. 1.9e-72;		
Matches 272; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0

QY	899	ATTTCAAAAGAAATCTAAAAAGGAAAGCAAAACAAAGAAATCTGAACAGAAAGTGGAAA	958
Db	54	ACTTTTAAAGGAATCTGAAAAAGGAAAGCAAAACAAAGAAATCTGAACAGAAAGTGGAAA	113
QY	959	TGAATACTCCGGCGACAAATTCAGATGATGAACAAAAGACACAGACCGCTTCTCTCCATCC	1018
Db	114	TGAAATCTCCGGCGACAAATTCAGATGATGAACAAAAGACACAGAAACCGCTTCTCTCTCC	173
QY	1019	CAAACTCTGTGAATCCAGTATCCTGTGAACAAAGAAACAGTACTGTGGCCCTCCCTACACA	1078
Db	174	CAAACTCTGTGAATCCAGTATCCTGTGAACAAAGAAACAGTACTGTGGCCCTCCCTACACA	233
QY	1079	GCGGGCTGATCAGTCACCCACAGCGCCAGCAAGTCATCCACAGGCTTCTTTAAATCTGGCC	1138

Page 3

LOCATION: 1..1310
SEQUENCE DESCRIPTION: SEQ ID NO: 13;

GENERAL INFORMATION:
APPLICANT: Kirschbaum, Bernd
Berglund, Erick

QY 172 PSDLSAVSIALTPQSGNCTLCVATWKSILKARKSATVNLTVIRCPQDT 220
 Db 182 PGSESTLTLTPFSHDDGATFVCRARSAQLPTGRDPAITLSLQYPPREV 230

RESULT 14

US-10-176-913-584
 ; Sequence 584, Application US/10176913
 ; Publication No. US2003002298A1
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C66
 CURRENT FILING DATE: 2002-06-20
 Prior Application removed - See file wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-176-913-584

Query Match

Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
 Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;

QY 14 SSGSNEVIEGPNARVILKSGARFNCVSGWKLMALSDMVLVSRPMEPIITNDREFT 73
 Db 20 AGSPHFLLQPEDLVLLGEARLPCALGAYWGLVMTKSGLAGGOR-----DLPG 71
 QY 74 SQRYDGGNFTS---EMIHNVEPSDSGNIRCSLONSRLHG-SAYLTV-----QVNGE 122
 Db 72 WSRWTISGNANQGHDLHIRPELEDEASYEQATQAGLSRPAQLHVLVPEAPQVLGG 131
 QY 123 LFPSTVNLVAENEPCEVTC-----LPSHWTWLPDISW-ELGLVSHSSY--FYPE-- 171
 Db 132 ---PSVSLVA--GVPA NLTCRSRGDARPT-----PELLMFRDGVLDGATFHQTLLKEGT 181
 QY 172 PSDLSAVSIALTPQSGNCTLCVATWKSILKARKSATVNLTVIRCPQDT 220
 Db 182 PGSESTLTLTPFSHDDGATFVCRARSAQLPTGRDPAITLSLQYPPREV 230

RESULT 15

US-10-180-552-584
 ; Sequence 584, Application US/10180552
 ; Publication No. US20030022300A1
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C153
 CURRENT APPLICATION NUMBER: US/10/180,552
 CURRENT FILING DATE: 2002-06-25
 Prior Application removed - See file wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 584
 LENGTH: 708
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-180-552-584

Query Match

Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
 Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;

QY 14 SSGSNEVIEGPNARVILKSGARFNCVSGWKLMALSDMVLVSRPMEPIITNDREFT 73
 Db 20 AGSPHFLLQPEDLVLLGEARLPCALGAYWGLVMTKSGLAGGOR-----DLPG 71
 QY 74 SQRYDGGNFTS---EMIHNVEPSDSGNIRCSLONSRLHG-SAYLTV-----QVNGE 122
 Db 72 WSRWTISGNANQGHDLHIRPELEDEASYEQATQAGLSRPAQLHVLVPEAPQVLGG 131
 QY 123 LFPSTVNLVAENEPCEVTC-----LPSHWTWLPDISW-ELGLVSHSSY--FYPE-- 171
 Db 132 ---PSVSLVA--GVPA NLTCRSRGDARPT-----PELLMFRDGVLDGATFHQTLLKEGT 181
 QY 172 PSDLSAVSIALTPQSGNCTLCVATWKSILKARKSATVNLTVIRCPQDT 220
 Db 182 PGSESTLTLTPFSHDDGATFVCRARSAQLPTGRDPAITLSLQYPPREV 230

Search completed: April 29, 2003, 20:17:26
 Job time : 47.0659 secs

```

RESULT 11
US-10-175-752-584
; Sequence 584, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;

QY 14 SSGSNEVIEGPONARVYKGSQARFNCYVSGWKLMALSDMYVLSVRPMEPIITNDRET 73
DB 20 AGSPHFHQDPEDLVLLGGEARLPCALGAYGVLQWTKSGIALGGQR-----DLPG 71
QY 74 SORYDQGNFTS---EMIHNVPSDSGNIRCSLONSRLHG-SAYLVY-----QWGE 122
DB 72 WSRWISGNANAGOHDLHIRPVELEDEASYEQATQAGLSRPAQLHVLVPPAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLVSHSSY--FVPE-- 171
DB 132 ---PSVSLVA--GYPANLTCRSRGDARPT-----PELWFRDGVLLDGAIFHOTLLKEGT 181
QY 172 PSDLOSASIALTPQSGNGLTCVATWKSLSKARKSATVNLTVIRCPQDT 220
DB 182 PGSEVSTLTPFSSHDDGATFVCARASQALPTGRDITATLSHQYPEEVT 230

RESULT 12
US-10-176-482-584
; Sequence 584, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm

```

```

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;

QY 14 SSGSNEVIEGPONARVYKGSQARFNCYVSGWKLMALSDMYVLSVRPMEPIITNDRET 73
DB 20 AGSPHFHQDPEDLVLLGGEARLPCALGAYGVLQWTKSGIALGGQR-----DLPG 71
QY 74 SORYDQGNFTS---EMIHNVPSDSGNIRCSLONSRLHG-SAYLVY-----QWGE 122
DB 72 WSRWISGNANAGOHDLHIRPVELEDEASYEQATQAGLSRPAQLHVLVPPAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLVSHSSY--FVPE-- 171
DB 132 ---PSVSLVA--GYPANLTCRSRGDARPT-----PELWFRDGVLLDGAIFHOTLLKEGT 181
QY 172 PSDLOSASIALTPQSGNGLTCVATWKSLSKARKSATVNLTVIRCPQDT 220
DB 182 PGSEVSTLTPFSSHDDGATFVCARASQALPTGRDITATLSHQYPEEVT 230

RESULT 13
US-10-176-757-584
; Sequence 584, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;

QY 14 SSGSNEVIEGPONARVYKGSQARFNCYVSGWKLMALSDMYVLSVRPMEPIITNDRET 73
DB 20 AGSPHFHQDPEDLVLLGGEARLPCALGAYGVLQWTKSGIALGGQR-----DLPG 71
QY 74 SORYDQGNFTS---EMIHNVPSDSGNIRCSLONSRLHG-SAYLVY-----QWGE 122
DB 72 WSRWISGNANAGOHDLHIRPVELEDEASYEQATQAGLSRPAQLHVLVPPAPQVYLG 131
QY 123 LFIPSVNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLVSHSSY--FVPE-- 171
DB 132 ---PSVSLVA--GYPANLTCRSRGDARPT-----PELWFRDGVLLDGAIFHOTLLKEGT 181

```

```
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;
```

```
QY 14 SSGNEVIEGPNARVLKSGQARFNCYVSGWKLIMALSDMVLSVRPEPIITNDRET 73
DB 20 AGSPHFHQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLAGGQR-----DLPG 71
QY 74 SQRVDGNGFTS---EMITHNVEPSDSGNIRCSLQNSRLHG-SAVLYV-----QYMG 122
DB 72 WSRWISGNAANGQHDHLIRPVELEDEASVCOATQAGLSRPAQLHVLVPPAPQVLGG 131
QY 123 LFIPSVNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLVSHSSSY--FVPE-- 171
DB 132 ---PSVSLVA--GVPANLTCRSRGDARPT-----PELLMFRDGVLLDGAFTFHQLKEGT 181
QY 172 PSDLOSASVILALTPQSNGLTLCVATWKSLSKARSAATVNLTVIRCPDPT 220
DB 182 PGVESTLTLLFPSSHDDGATFVCRARSQALPTGRDTAITLSLQYPPPEVT 230
```

```
RESULT 9
US-10-173-706-584
; Sequence 584, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
```

```
; ORGANISM: Homo Sapien
US-10-173-706-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;
```

```
QY 14 SSGNEVIEGPNARVLKSGQARFNCYVSGWKLIMALSDMVLSVRPEPIITNDRET 73
DB 20 AGSPHFHQPEDLVLLGGEARLPCALGAYWGLVQWTKSGLAGGQR-----DLPG 71
QY 74 SQRVDGNGFTS---EMITHNVEPSDSGNIRCSLQNSRLHG-SAVLYV-----QYMG 122
DB 72 WSRWISGNAANGQHDHLIRPVELEDEASVCOATQAGLSRPAQLHVLVPPAPQVLGG 131
QY 123 LFIPSVNLVVAENPECEVTC-----LPSHWTWLPDISW-ELGLVSHSSSY--FVPE-- 171
DB 132 ---PSVSLVA--GVPANLTCRSRGDARPT-----PELLMFRDGVLLDGAFTFHQLKEGT 181
QY 172 PSDLOSASVILALTPQSNGLTLCVATWKSLSKARSAATVNLTVIRCPDPT 220
DB 182 PGVESTLTLLFPSSHDDGATFVCRARSQALPTGRDTAITLSLQYPPPEVT 230
```

```
RESULT 10
US-10-175-738-584
; Sequence 584, Application US/10175738
; Publication No. US2003002294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-584

Query Match
Best Local Similarity 23.1%; Score 128; DB 9; Length 708;
Matches 53; Conservative 41; Mismatches 95; Indels 40; Gaps 11;
```

US-10-175-737-584
; Sequence 584, Application US/10175737

Db 362 ILNOR-----ORGERKAPENOEEREAELN----- 390
Qy 325 OKTTDASLPKSCSSDPEORNSCGPP 353
Db 391 -----QSEPEPAGESSSTGP 405

RESULT 2

US-10-184-644-559
; Sequence 559, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Goddowskl, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 559
LENGTH: 2473
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-559

Query Match
Best Local Similarity 29.5%; Score 138.5; DB 9; Length 2473;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

Qy 183 ALTPQSGTLTCVATWKSARKSATVNLVIRCPDGTGGGINIPGLSLPSLGFSLPT 242
Db 2274 AATGAGTTTCAATTAATTAATTAATGTTTC----- 2307

Qy 243 WGVVGLAGTMLTPCTLTIRCCCR-RRCGCNCCRCFCFC 286
Db 2308 -----ATTCCATCGCCACCCACCCGCCGCCGCCACCC 2344

RESULT 3

US-10-184-634-559
; Sequence 559, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Goddowskl, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 559

LENGTH: 2473
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-559

Query Match
Best Local Similarity 29.5%; Score 138.5; DB 9; Length 2473;
Matches 31; Conservative 3; Mismatches 36; Indels 35; Gaps 2;

Qy 183 ALTPQSGTLTCVATWKSARKSATVNLVIRCPDGTGGGINIPGLSLPSLGFSLPT 242
Db 2274 AATGAGTTTCAATTAATTAATTAATGTTTC----- 2307

Qy 243 WGVVGLAGTMLTPCTLTIRCCCR-RRCGCNCCRCFCFC 286
Db 2308 -----ATTCCATCGCCACCCACCCGCCGCCGCCACCC 2344

RESULT 4

US-10-016-283-33
; Sequence 33, Application US/10016283
; Patent No. US20020164702A1
; GENERAL INFORMATION:

APPLICANT: Valenzuela et al., David M.
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
FILE REFERENCE: REG195-B-PCT-US
CURRENT APPLICATION NUMBER: US/10/016,283
CURRENT FILING DATE: 2001-11-30
Prior Application NUMBER: US/09/077,955A
Prior Filing DATE: 1998-09-10
Prior Application NUMBER: PCT/US96/20696
Prior Filing DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 33
LENGTH: 869
TYPE: PRT
ORGANISM: Homo sapiens
US-10-016-283-33

Query Match
Best Local Similarity 20.8%; Score 132.5; DB 9; Length 869;
Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;

Qy 4 GARNRDPGGSGGNEVIEGPQARVLKSGQARNCVSGWK-LIMVALSDMYLSVRP 62
Db 113 GALQYKMP-----KITRPINVKIIEGLKAVLPCTTMGNPKPSVSMIKGD----- 158

Qy 63 MEPTTNDRTSQRVDGNTSEMIITHNVPDSQNRCSLQSRKLSGNY-LTVQYMG 121
Db 159 -SPLENSRIAVLE-----SGSLRIHNQKEDAGQRCVAKNSL--GRAYSKVYLEV 208

Qy 122 ELFTPSVNLVVAENEP-----CEVTCLPSHTWLPDISW-ELGLVSHSYVFPPEP 172
Db 209 EVFARILFAPSHNVTRGSFVTLHCTATGTP-----VPTIWIENGNAVSSGSIQESYKD 263

Qy 173 SDLOSAYSILALTPQSGTLTCVATWKSARKSATVNLVIRCPDGTGG----- 223
Db 264 RVIDSRQLFTKRP---GLTYCIATNKHGEKFSYAKAAATSIENWPKPKDKMNGYCAQY 320

Qy 224 -----INIPGVLSLPSLGSFSL-----PTWKGVLGLAGTMLTPCTLTIRCC 267
Db 321 RGEVCNNAVLAARDALVFNITSYADPEAEOLLVHTAMNEL-----KVSPVCPAPAEAL 373

Qy 268 CRRRCGCCN-----CCCRCC-----FCRRRRGRFRIQKSEKTKETKETE 312
Db 374 LCNHIFQCSGVVPTPTPCREYCLAVKELFKAKE---WLVWEKTHGLRSEHML 429

Qy 313 SGENSGYNSDEQKTTDTASLPKSCSSDPEORNSCGPPHARQRP-----PRPASHP 368
Db 430 SVPEGSKLPSNHMPPTACARLP-----HLDYKNKENTLTPPP--MTSSKPEVDLPNLSSS 482

Qy 369 QASFNLA 375

Result No.	Score	Query Match	Length	DB	ID	Description
1	140	6.7	405	1	US-08-755-235-4	Sequence 4, Appli
2	138.5	6.6	2473	9	US-10-184-644-559	Sequence 559, App
3	138.5	6.6	2473	9	US-10-184-634-559	Sequence 559, App
4	132.5	6.3	869	9	US-10-016-283-33	Sequence 33, Appl
5	131.5	6.3	869	10	US-09-817-487A-2	Sequence 2, Appl
6	128	6.1	708	9	US-10-174-590-584	Sequence 584, App
7	128	6.1	708	9	US-10-176-758-584	Sequence 584, App
8	128	6.1	708	9	US-10-173-737-584	Sequence 584, App
9	128	6.1	708	9	US-10-175-706-584	Sequence 584, App
10	128	6.1	708	9	US-10-175-738-584	Sequence 584, App
11	128	6.1	708	9	US-10-175-752-584	Sequence 584, App
12	128	6.1	708	9	US-10-176-482-584	Sequence 584, App
13	128	6.1	708	9	US-10-176-757-584	Sequence 584, App
14	128	6.1	708	9	US-10-176-913-584	Sequence 584, App
15	128	6.1	708	9	US-10-180-552-584	Sequence 584, App
16	128	6.1	708	9	US-10-180-557-584	Sequence 584, App
17	128	6.1	708	9	US-10-173-700-584	Sequence 584, App
18	128	6.1	708	9	US-10-174-572-584	Sequence 584, App
19	128	6.1	708	9	US-10-174-579-584	Sequence 584, App

20	128	6.1	708	9	US-10-174-583-554	Sequence 584, App
21	128	6.1	708	9	US-10-174-588-554	Sequence 584, App
22	128	6.1	708	9	US-10-175-723-554	Sequence 584, App
23	128	6.1	708	9	US-10-175-740-554	Sequence 584, App
24	128	6.1	708	9	US-10-175-743-554	Sequence 584, App
25	128	6.1	708	9	US-10-176-488-554	Sequence 584, App
26	128	6.1	708	9	US-10-176-492-554	Sequence 584, App
27	128	6.1	708	9	US-10-176-747-554	Sequence 584, App
28	128	6.1	708	9	US-10-176-750-554	Sequence 584, App
29	128	6.1	708	9	US-10-176-985-554	Sequence 584, App
30	128	6.1	708	9	US-10-176-987-554	Sequence 584, App
31	128	6.1	708	9	US-10-176-991-554	Sequence 584, App
32	128	6.1	708	9	US-10-176-992-554	Sequence 584, App
33	128	6.1	708	9	US-10-176-993-554	Sequence 584, App
34	128	6.1	708	9	US-10-184-658-554	Sequence 584, App
35	128	6.1	708	9	US-10-173-695-554	Sequence 584, App
36	128	6.1	708	9	US-10-173-697-554	Sequence 584, App
37	128	6.1	708	9	US-10-173-705-554	Sequence 584, App
38	128	6.1	708	9	US-10-174-576-554	Sequence 584, App
39	128	6.1	708	9	US-10-174-585-554	Sequence 584, App
40	128	6.1	708	9	US-10-174-586-554	Sequence 584, App
41	128	6.1	708	9	US-10-175-747-554	Sequence 584, App
42	128	6.1	708	9	US-10-176-481-554	Sequence 584, App
43	128	6.1	708	9	US-10-176-485-554	Sequence 584, App
44	128	6.1	708	9	US-10-176-487-554	Sequence 584, App
45	128	6.1	708	9	US-10-176-493-554	Sequence 584, App

```

Sequence 4 Application US/08755235
Publication No. US20030059423A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Schmidt, Ann Marie
APPLICANT: Wu, Jun
TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
FILE REFERENCE: 0575/50159
CURRENT APPLICATION NUMBER: US/08/755,235
CURRENT FILING DATE: 1996-11-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 405
TYPE: PRT
ORGANISM: Human
US-08-755-235-4

```

Matches 78; Conservative 34; Mismatches 93; Indels 124; Gaps 16

```

QY 64 EPITLTDRTS-----QXRYDGGNFT--SEMIITHNVPESDSCNIR----CSLQNSRLHG 112
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 162 KPLVPEKESVSKEDQTRKHPETGLTQSELM-----VTPANGDPRTESSCSPELPYHR 218
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 113 AYLTVQVMGELFIP-----SVNLVAENEP-----CEVTLSPHMTWLPDIS 154
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 219 ALRTAFIQRWMEPRVPLEEYOLVV---EPKGAVAVGGITVLTCEPPAPPS-----PQIH 270
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 155 WETGLVSHSSYYFPEPESDQSAVSIALTPQSNGLTCVATWKSIAKRS-ATVNTV 213
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 271 WMKD-----GVPLPLPSPLILPELGPQDQGYSCVATHSHGQESRAVYSI 324
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 214 IRPQDTGGGINIPVLSLSPLSGFSILPFWKYGKGLAGMMLT-----PTCLTITRC 266
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 322 IP-PGEG-----PTGSGVGGSGGLIALALGILAGLTAALLGV 361
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 267 CCCRRCGCCGCCRCPCFCRCRRKRGFIQPKKSEKRT--NKETTESGNGNSGYNSDE 324
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```


C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro
 C:Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco
 structural protein

F:24752-25008/Domain: protein kinase homology <KIN>
 F:84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,
 98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1
 tatus predicted
 F:16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18
 ,21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248
 F:26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 5.5%; Score 114.5; DB 1; Length 26926;

Best Local Similarity 22.1%; Pred. No. 45;

Matches 62; Conservative 31; Mismatches 87; Indels 101; Gaps 12;

QY 3 AGAMENDPPGSGSGNEVIEGPQ-----NARV-----LKGQARENCIVS 42

Db 16422 AGVEHADVPGPPIVEEKLAPDIDLLELRKIINIRAGSLRLEVPKGRP-----TPE 16476

QY 43 QGRKLIMWALSDMYVLSVRMEPIITNDRTSQRYDQGNFTSEMIHNHNEPSDSGNIRC 102

Db 16477 VKMGKVDGEIRDAIIDVT-----SSFTS-LVLDNVRNRYDSGKRYTL 16516

QY 103 SLONSRLHGSAYLTVQYMGELFIPSVNLVVAENEPCEVTCLPSSHWTWLPDI----- 153

Db 16517 TLENSSGTKSAFYVRYL-DTPSPVNLKYTEIKDSVST-----TWEPPLDGGSKIKN 16570

QY 154 -----SWEIGLLVSHSSYTF-----VPEPSDLQSAVST 181

Db 16571 YIVEKREATIRKSYAAVYTNCHKNSWKIDQLQEGCSYFRVTAENEXGIGLPAQTADPIKY 16630

QY 182 LALTPQSGTLTCVATWKSILKARKSATVNLTVIRCPDQTGG 222

Db 16631 -AEVPQPPGKITVDV-----TRNSVSLSWTK---PEHDGG 16662

Search completed: April 28, 2003, 21:09:09
 Job time : 40.0659 secs

Db 145 TPEVKGKVDGDIRDAIIDVTSSFTS-LVLDNVRNDSGKYLTLLENSSGTKSAFYVR 203
 QY 119 VMGELFIPSVNLVVAENECVETCLPSHTWLPDI----- 153
 Db 204 VL-DTSPPVNKLKVEITIDSVI-----TWEPPLLDGSKIKNIYVEKREKRSYAAV 257
 QY 154 -----SMELGLVSHSSYFVPEPSDLSAVSILALT-----POSNGTLTCVAT 197
 Db 258 VVNHCKNSMKIDQLGSCSYF-RVTAENEYGIAGAFADPIKVAEYVQPPGKITVDV 316
 QY 198 WSKLRKRSATVNLVYRCPDTCG 222
 Db 317 -----TRNSYSLSWTK---PEHDGG 333

RESULT 14

differentiation antigen - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: I49583

R:Law, C.

J. Immunol. 151, 175-187, 1993

A:Title: Organization of the murine Cd22 locus. Mapping to chromosome 7 and characteriza

A:Reference number: I49583; MUID:93315834; PMID:8100843

A:Accession: I49583

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-862 <RES>

A:Cross-references: GB:I49583; MUID:9348965; PIDN:AAA02562.1; PID:9348966

C:Genetics: CD22

Query Match

Best Local Similarity 19.1%; Pred. No. 1.1; DB 2; Length 862;

Matches 86; Conservative 36; Mismatches 126; Indels 203; Gaps 18;

QY 10 DPFGSGGNEVIEIPONARVLKSGQARFN-----CTVSQGKWL-----IMMA 51
 Db 472 NPGSGS-----VLKPGVLRICKVTWDSMPVSCACNHCMSALPILNVHVA 519
 QY 52 LSDMYLVSRMEPIITNDRTFSQ-----RIDQGNFTSE---MIHNVES 95
 Db 520 PRDVKVLKVPASEIRAGQVLLQCDPABESNPAEVRFFPKNGSLVQEGRYLSEFGSVSE 579
 QY 96 DSGNIRGSLNRLHSGAVYTVGMGELFIPSVNLV-----VAENEC 138
 Db 580 DSGYINCMVNN-----IGETLSQAWNLVLYAPRLRVISIPGDHVMGKKA 627
 QY 139 EYTCLP-----SHWTWLPDISWELGLVSHSSYFVPEPSDLSAVSILALT---QSN 189
 Db 628 TLSCESDANPIQIYTW-----FDSSQGLDHSQGLKRLREPLEVQHT 669
 QY 190 GTLTCAVTWMSLKARSAVNLVYRCPDTCGGINIPVGLSLPGLGSLPVGKVG 249
 Db 670 GSYRCKCT-NGIGTGESPESTLVYSPETIG-----KRAVAG 706
 QY 250 LAGTMLTPTCTLIRCCCRRCGCCGCCRCFCCKRRKGRFRI-----FOK 298
 Db 707 LGFCLIT---CLAI-----WGKTIQKKKKQNSQGLDE 738
 QY 299 KSE-----EKTNKETETESGNSG-----YNSDEKTTDT 330
 Db 739 NSSQGSFFVNRKKARPTPLSEGGQSCYVPAAMDITVSAIILRFPSMDHNMAGDAGTPAT 798
 QY 331 ASLPKSCS-----SPEDQNSC 350
 Db 799 QAPPNNSDVYTSYVIOKRPMDGYENVNPS 829

RESULT 15

I38344 titin, cardiac muscle [validated] - human

N:Alternate names: connectin
 N:Contents: serine/threonine-specific protein kinase (EC 2.7.1.-)
 C:Species: Homo sapiens (hmn)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Sep-2000
 C:Accession: I38344; I38345; S20898; S20897; S20899; S6365; S37393
 R:Label: S.; Kolmerer, B.
 Science 270, 293-296, 1995
 A:Title: Tittins: giant proteins in charge of muscle ultrastructure and elasticity.
 A:Reference number: A57430; MUID:96026330; PMID:7569978
 A:Accession: I38344
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/
 A:Molecule type: mRNA
 A:Residues: 1-26926 <LAB1>
 A:Cross-references: EMBL:X90568; MUID:91017424; PID:91017425
 R:Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.
 Biochemistry 34, 553-561, 1995
 A:Title: Dissecting titin into its structural motifs: identification of an alpha-hel
 A:Reference number: I38345; MUID:95119041; PMID:7819249
 A:Accession: I38345
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1977-2014 <MUS>
 A:Cross-references: EMBL:X83270; MUID:9602579; PIDN:CAA58243.1; PID:9602580
 A>Note: conformation and properties are reported for a synthetic peptide correspondin
 R:Label: S.; Gautel, M.; Lakey, A.; Trinick, J.
 EMBO J. 11, 1711-1716, 1992
 A:Title: Towards a molecular understanding of titin.
 A:Reference number: S20897; MUID:92258380; PMID:1582406
 A:Accession: S20898
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 13597-14200, 'I', 14202-14696 <LAB2>
 A:Cross-references: EMBL:X64698; MUID:937192; PIDN:CAA5939.1; PID:937193
 A:Accession: S20897
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 16330-16382, 'S', 16384-16756, 'F', 16758-16860 <LAB3>
 A:Cross-references: EMBL:X64699; MUID:937190; PIDN:CAA5940.1; PID:937191
 A:Accession: S20899
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 2248
 A:Cross-references: EMBL:X64697; MUID:937190; PIDN:CAA5938.1; PID:937195
 R:Kolmerer, B.; Olivieri, N.; Wilt, C.C.; Hermann, B.G.; Labelle, S.
 J. Mol. Biol. 256, 556-563, 1996
 A:Title: Genomic organization of M line titin and its tissue-specific expression in t
 A:Reference number: S6365; MUID:96177761; PMID:8604138
 A:Accession: S6365
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 26729-26825 <KOL>
 A:Cross-references: EMBL:X92412; MUID:91236761
 R:Gautel, M.; Leonard, K.; Labelle, S.
 EMBO J. 12, 3827-3834, 1993
 A:Title: Phosphorylation of KSP motifs in the C-terminal region of titin in different
 A:Reference number: S37393; MUID:94008990; PMID:8404852
 A:Accession: S37393
 A:Molecule type: mRNA
 A:Residues: 26831-26926 <GAU>
 R:Improta, S.; Politou, A.S.; Pastore, A.
 submitted to the Brookhaven Protein Data Bank, February 1996
 A:Reference number: A6736; PDB:1TIT
 A:Contents: annotation: conformation by (1)H-NMR, residues 5253-5341
 R:Pfuhl, M.; Pastore, A.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A:Reference number: A66201; PDB:1NCT
 A:Contents: annotation: conformation by (1)H-NMR, residues 'S', 26059-26155
 C:Genetics: CD22
 A:Gene: GDB:TTN
 A:Cross-references: GDB:127867; OMIM:188840
 A:Map position: 2q31-2q32
 A:Description: structural protein forming filaments in striated muscle

QY 120 MGELE-IPSVNLVAENE---PCEVLCIPSHWTWLPDISN-ELGLLVSHSSYYFVPEPSD 174

63 MEPT-----ITNDRTSORXDQGNFTSEMIIHNVEPDSGIRCSLONSRLHGSAYLTVQ 118

```

Db 274 LRGEVLT-----QLRSKYSYLLGG--SNLLISNYDDSDSGMTVCVYTKNEN 318
QY 109 LHSAYTLVQVMEGLFIPSVNLVVAENPECEVYCLPSHMTWLPDISW-ELGLVSHSSY 167
Db 319 ISASAEITLVPPWFLNHPNSLIAYESMDIEFFECTVSGKP-VPTVNMKNKGVDVYPSDYF 377
QY 168 FVPEPSDLSQASVILALTPQSNGLTCVATWKSILKARKSATVNLVIRCPDPTGGGINIP 227
Db 378 QIVGGSNLR-----ILGVKSDSEGFYQCAENENAGNQTSAQLVPRPAIPSSS----- 426
QY 228 GVLSLIP 234
Db 427 -VLPSAP 432

RESULT 2
PCT-US94-05277-2
; Sequence 2, Application PC/TUS9405277
; GENERAL INFORMATION:
; APPLICANT: Bruskin, Arthur
; APPLICANT: Jarosz, David E.
; APPLICANT: Johnson, Karen
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Antibodies specific for DCC Gene Product
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05277
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42709
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-05277-2

Query Match 6.5%; Score 136; DB 5; Length 1447;
Best Local Similarity 25.1%; Pred. No. 0.0018;
Matches 62; Conservative 34; Mismatches 101; Indels 50; Gaps 10;
QY 9 RDPGSGSGNEV-----IGPONARVLKGSQARFNTVSQGW--KLIMW 50
Db 215 RNPASSRTGNAEYRIISDPELHROLYFLQRPNSVAIEGKDAVLECCVS-GTIPPSFTW 273
QY 51 ALSDMVTVSRMPEITITNDRFTSQRYDOGNGFTSEMIINHPSPSDSGNTRC--SLQNSR 108
Db 274 LRGEVLT-----QLRSKYSYLLGG--SNLLISNYDDSDSGMTVCVYTKNEN 318
QY 109 LHSAYTLVQVMEGLFIPSVNLVVAENPECEVYCLPSHMTWLPDISW-ELGLVSHSSY 167

```

```

Db 319 ISASAEITLVPPWFLNHPNSLIAYESMDIEFFECTVSGKP-VPTVNMKNKGVDVYPSDYF 377
QY 168 FVPEPSDLSQASVILALTPQSNGLTCVATWKSILKARKSATVNLVIRCPDPTGGGINIP 227
Db 378 QIVGGSNLR-----ILGVKSDSEGFYQCAENENAGNQTSAQLVPRPAIPSSS----- 426
QY 228 GVLSLIP 234
Db 427 -VLPSAP 432

RESULT 3
US-08-374-834-16
; Sequence 16, Application US/08374834
; Patent No. 5656473
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,834
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/095,658
; FILING DATE: 21-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Coibert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 190A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 869 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-374-834-16

Query Match 6.3%; Score 132.5; DB 1; Length 869;
Best Local Similarity 20.8%; Pred. No. 0.0018;
Matches 89; Conservative 55; Mismatches 178; Indels 105; Gaps 20;
QY 4 GAMENDPPGSGSGNEVIEGPNARVLKGSQARFNTVSQGW--KLIMW 62
Db 113 GAOVWKP-----KTRPPINVKITIBGLKAVLPCTTMGNKPSVSIKGD----- 158
QY 63 MPEITITNDRFTSQRYDOGNGFTSEMIINHPSPSDSGNTRC--SLQNSR 121
Db 159 -SPLRENSKIAVLE-----SGSLRIHNVQKEDAGQRCVAKNKL--GTAYSKVVKLEV 208
QY 122 ELFIPSVNLVVAENEP-----CEVTCPSHMTWLPDISW-ELGLVSHSSYTFVEP 172
Db 209 EVFARILRAPESHNVTFGSGFTVILHCTATGIP-----VPTIWIENAGNAVSSGSIQESYKD 263
QY 173 SLDQASVSLALTPQSNGLTCVAT-----WKSILKARKSATVNLVIRCPDPTGGG----- 223

```

```

1      RESULT 1
2      : US-09-041-886-25
3      : Sequence 25, Application US/09041866
4      : Patent No. 6235872
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Bredesen, Dale E.
8      : APPLICANT: Rabinzadeh, Sharoz
9      : TITLE OF INVENTION: Prapoptotic Peptides, Dependence
10     : TITLE OF INVENTION: Polypeptides and Methods of use
11     : NUMBER OF SEQUENCES: 72
12     :
13     : CORRESPONDENCE ADDRESS:
14     : ADDRESSEE: Campbell & Flores LLP
15     : STREET: 4370 La Jolla Village Drive, Suite 700
16     : CITY: San Diego
17     : STATE: California
18     : COUNTRY: United States
19     : ZIP: 92122
20     :
21     : COMPUTER READABLE FORM:
22     : MEDIUM TYPE: Floppy disk
23     : COMPUTER: IBM PC compatible
24     : OPERATING SYSTEM: PC-DOS/MS-DOS
25     : SOFTWARE: Patentin Release #1.0, Version #1.25
26     :
27     : CURRENT APPLICATION DATA:
28     : APPLICATION NUMBER: US/09/041.886
29     : FILING DATE:
30     : CLASSIFICATION:
31     : ATTORNEY/AGENT INFORMATION:
32     : NAME: Campbell, Cathryn A.
33     : REGISTRATION NUMBER: 31,815
34     : REFERENCE/DOCKET NUMBER: P-LJ 2626
35     : TELECOMMUNICATION INFORMATION:
36     : TELEPHONE: (619) 535-9001
37     : TELEFAX: (619) 535-8949
38     : INFORMATION FOR SEQ ID NO: 25:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 1447 amino acids
41     : TYPE: amino acid
42     : TOPOLOGY: linear
43     : MOLECULE TYPE: protein
44     :
45     : US-09-041-886-25
46
47     Query Match 6.5%; Score 136; DB 4; Length 1447;
48     Best Local Similarity 25.1%; Pred. No. 0.0018;
49     Matches 62; Conservative 34; Mismatches 101; Indels 50; Gaps
50
51     QY 9 RDPGSGSGSNEY-----IEGPONARVLKGSQARFNCVTYSQGM--KLIMW 50
52     | | | | | | | | | | | | | | | | | | | | | | | | | | | |
53     Db 215 RNASSRTGNEAEVRLIDPGLHQILFLDLPSPSVYVAIEBKDAVLECYVS-GYPPSPSTW 273
54     Y 51 ALSDMVIVLVRPEEPIITNDRFTSQRYDQGNFTSEMIINHPSPDSGNIRC--SLQNSR 108

```


MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
IMMEDIATE SOURCE:
CLONE: lambda yPt
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626
US-08-258-442-12

Query Match 3.4%; Score 42.6; DB 1; Length 3292;
Best Local Similarity 64.9%; Pred. No. 0.033;
Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 869 ATACGCTGCTGCTGCGCCGCTGTTGAGCTGCAACGCTGCGCGTGTGT 928
DB 818 AGATGATGATGTTGTTGTAAGTTGTGTGTTGCAACTGCTGCTGTAGC 759

OY 929 TTCTGCTGTAAGAAAAAGAGATTTCGTAATTCAT 965
DB 758 TCCTGCTGTGCTGCAACTGCTGTGTGTGTAATT 722

RESULT 7
US-08-328-809-7/c

Sequence 7, Application US/08328809
Patent No. 5705334

GENERAL INFORMATION:

APPLICANT: Lippard, Stephen J.

APPLICANT: Essigmann, John M.

APPLICANT: Donahue, Brian A.

APPLICANT: Toney, Jeffrey H.

APPLICANT: Bruhn, Suzanne L.

APPLICANT: Pili, Pieter M.

APPLICANT: Brown, Steven

APPLICANT: Kelleff, Patil

TITLE OF INVENTION: Uses for DNA Structure-Specific

TITLE OF INVENTION: Recognition Proteins

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator, Testa, Hurwitz & Thibault

STREET: 53 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,809

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fenton, Gillian M.

REGISTRATION NUMBER: 36,508

REFERENCE/DOCKET NUMBER: MIT-023 (5473/24)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-248-7000

TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 3292 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Saccharomyces cerevisiae

IMMEDIATE SOURCE:

CLONE: lambda yPt

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626
US-08-328-809-7

Query Match 3.4%; Score 42.6; DB 1; Length 3292;
Best Local Similarity 64.9%; Pred. No. 0.033;
Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 869 ATACGCTGCTGCTGCGCCGCTGTTGAGCTGCAACGCTGCGCGTGTGT 928
DB 818 AGATGATGATGTTGTTGTAAGTTGTGTGTTGCAACTGCTGCTGTAGC 759

OY 929 TTCTGCTGTAAGAAAAAGAGATTTCGTAATTCAT 965
DB 758 TCCTGCTGTGCTGCAACTGCTGTGTGTGTAATT 722

RESULT 8
PCT-US92-11107-12/c

Sequence 12, Application PC/TUS9211107
GENERAL INFORMATION:

APPLICANT: Donahue, Brian A.

APPLICANT: Toney, Jeffrey H.

APPLICANT: Bruhn, Suzanne L.

APPLICANT: Pili, Pieter M.

APPLICANT: Brown, Steven

APPLICANT: Kelleff, Patil

APPLICANT: Essigmann, John M.

APPLICANT: Lippard, Stephen J.

TITLE OF INVENTION: DNA Structure Specific Recognition

TITLE OF INVENTION: Protein and Uses Therefor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: 2 Millilla Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/11107

FILING DATE: 19921218

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/539,906

FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: MIT-4787AAA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 3292 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Saccharomyces cerevisiae

IMMEDIATE SOURCE:

CLONE: lambda yPt

NAME/KEY: CDS

LOCATION: 1..1626

TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1116
US-09-041-886-12

Query Match 3.5%; Score 42.8; DB 4; Length 1776;
Best Local Similarity 61.8%; Pred. No. 0.021;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 873 GCTGCTGCTGCTGCGCGCGCTGTTGTTGCTGCACTGCTGCTGCTGTTTCT 932
DB 975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTT 916
QY 933 GCTGTAGAGAAAGAGAGATTGCTATTCATTCAATTCAAGAAATCTGAA 982
DB 915 GCTGCTGTTTTCAGAGTAGGCTTCTGCTCTTCGACAGCTCTTGAA 866

RESULT 5

US-07-814-964-12/C
Sequence 12, Application US/07814964
Patent No. 5359047

GENERAL INFORMATION:

APPLICANT: Donahue, Brian A.
APPLICANT: Toney, Jeffrey H.
APPLICANT: Bruhn, Suzanne L.
APPLICANT: Pil, Pieter M.
APPLICANT: Brown, Steven
APPLICANT: Kellelt, Patli
APPLICANT: Essigmann, John M.
APPLICANT: Lippard, Stephen J.
TITLE OF INVENTION: DNA Structure Specific Recognition
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: 2 Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/814,964
FILING DATE: 19911226
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539,906
FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4787AAA
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 3292 base pairs
TYPE: NUCLEIC ACID

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
IMMEDIATE SOURCE:
CLONE: lambda yPt
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626
US-07-814-964-12

Query Match 3.4%; Score 42.6; DB 1; Length 3292;
Best Local Similarity 64.9%; Pred. No. 0.033;
Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 869 ATACGCTGCTGCTGCGCGCGCTGTTGTTGCTGCACTGCTGCGTGTGTGT 928
DB 818 AGATGATGATGATGTTGTTGAAGTTGTTGTTGTTGCACTGCTGTTGTAAC 759
QY 929 TTCTGCTAGAGAAAGAGATTGCTATTCAT 965
DB 758 TGCTGCTGTTGCTGCACTGCTGTTGTTGTAAT 722

RESULT 6

US-08-258-442-12/C
Sequence 12, Application US/08258442
Patent No. 5670621

GENERAL INFORMATION:

APPLICANT: Donahue, Brian A.
APPLICANT: Toney, Jeffrey H.
APPLICANT: Bruhn, Suzanne L.
APPLICANT: Pil, Pieter M.
APPLICANT: Brown, Steven
APPLICANT: Kellelt, Patli
APPLICANT: Essigmann, John M.
APPLICANT: Lippard, Stephen J.
TITLE OF INVENTION: DNA Structure Specific Recognition
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: 2 Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,442
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539,906
FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4787AAA
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 3292 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,927B
FILING DATE: 21-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP H6-251600
FILING DATE: 21-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: ATH95-01A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1115
US-08-531-927B-1

Query Match 3.5%; Score 42.8; DB 2; Length 1776;
Best Local Similarity 61.8%; Pred. No. 0.021; Indels 0; Gaps 0.
Matches 68; Conservative 0; Mismatches 42;

OY 873 GCTGCGTCTGCTCCGCCCGCTGTTGTGGGCTGCACTGCTGCGCGCTTTGTTCT 932
Db 975 GCTGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTT 916
OY 933 GCTGTGAAGAAAGAGAGATTTCGTATTCATTTCAAAAGAAATTCGAA 992
Db 915 GCTGCGTGTTTTTCAAAGTAGGCTCTCGTCTTCCGAAAGCTCTTCGAA 866

RESULT 4
US-09-041-886-12/c
Sequence 12, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Bredesen, Dale E.
APPLICANT: Rabilzadeh, Sharroz
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
Zip: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003 ; Search time 51.9118 Seconds

7325.486 Million cell updates/sec

Title:	US-09-729-264-5
Perfect score:	1340

Sequence: 1 aggtgtgtgagtcagccaaca.....gtaatacaactgtagtatag 1240

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq Length: 0
Maximum DB seq Length: 2000000000
```

Post-processor: Match 09

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

```

1: /cgn2_6/ptodata1/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata1/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata1/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata1/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata1/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata1/1/ina/backfiles1.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	51.8	4.2	397	3	US-09-253-691-3..	Sequence 3, Appl1
C 2	48.2	3.9	325	2	US-08-531-927B-3..	Sequence 3, Appl1
C 3	42.8	3.5	1776	2	US-08-531-927B-1..	Sequence 1, Appl1
C 4	42.8	3.5	1776	4	US-09-041-886-12	Sequence 12, Appl1
C 5	42.6	3.4	3292	1	US-07-814-964-12	Sequence 12, Appl1
C 6	42.6	3.4	3292	1	US-08-258-442-12	Sequence 12, Appl1
C 7	42.6	3.4	3292	1	US-08-328-809-7	Sequence 7, Appl1
C 8	42.6	3.4	3292	5	PCT-US92-11107-18	Sequence 12, Appl1
C 9	42.6	3.4	40000	4	US-09-780-049-12	Sequence 12, Appl1
C 10	42.2	3.4	2674	4	US-09-817-180-1..	Sequence 1, Appl1
C 11	41.6	3.4	78	4	US-09-043-303-12	Sequence 12, Appl1
C 12	41.6	3.4	203	4	US-09-043-303-7	Sequence 7, Appl1
C 13	41.2	3.3	379	1	US-08-145-617-5	Sequence 5, Appl1
C 14	41.1	3.3	3771	1	US-08-185-432-3	Sequence 3, Appl1
C 15	41.1	3.3	3771	1	US-08-185-432-23	Sequence 23, Appl1
C 16	41.1	3.3	5053	1	US-08-185-432-1	Sequence 1, Appl1
C 17	40.8	3.3	6177	4	US-08-479-913E-1	Sequence 1, Appl1
C 18	40.8	3.3	7257	4	US-09-091-042A-1	Sequence 1, Appl1
C 19	40.4	3.3	4864	6	5496550-5	Patent No. 5496555
C 20	40.4	3.3	10348	2	US-08-457-273B-41	Sequence 41, Appl1
C 21	40.4	3.3	10348	3	US-08-556-419-13	Sequence 13, Appl1
C 22	40.4	3.3	10348	4	US-09-041-886-14	Sequence 14, Appl1
C 23	40.4	3.3	10366	1	US-08-246-982A-5	Sequence 5, Appl1
C 24	40.4	3.3	10366	1	US-08-453-265-5	Sequence 5, Appl1
C 25	40	3.2	75	4	US-09-043-303-10	Sequence 10, Appl1
C 26	40	3.2	78	4	US-09-043-303-11	Sequence 11, Appl1
C 27	40	3.2	154	1	US-08-469-302B-6	Sequence 6, Appl1

C 28	40	3.2	154	2	US-08-267-803B-1	Sequence 6, Appl 1
C 29	40	3.2	165	4	US-09-043-303-17	Sequence 17, Appl 1
C 30	40	3.2	168	1	US-08-469-802B-4	Sequence 4, Appl 1
C 31	40	3.2	168	2	US-08-267-803B-4	Sequence 4, Appl 1
C 32	40	3.2	171	1	US-08-469-802B-5	Sequence 5, Appl 1
C 33	40	3.2	171	2	US-08-267-803B-5	Sequence 5, Appl 1
C 34	40	3.2	193	1	US-08-469-802B-2	Sequence 2, Appl 1
C 35	40	3.2	195	2	US-08-267-803B-2	Sequence 2, Appl 1
C 36	40	3.2	234	1	US-08-469-802B-3	Sequence 3, Appl 1
C 37	40	3.2	234	2	US-08-267-803B-3	Sequence 3, Appl 1
C 38	40	3.2	270	4	US-09-146-054-8	Sequence 8, Appl 1
C 39	40	3.2	270	4	US-09-664-977A-8	Sequence 8, Appl 1
C 40	40	3.2	477	4	US-09-135-994-1	Sequence 1, Appl 1
C 41	40	3.2	506	1	US-08-469-802B-7	Sequence 7, Appl 1
C 42	40	3.2	506	2	US-08-267-803B-7	Sequence 7, Appl 1
C 43	40	3.2	623	4	US-09-043-303-5	Sequence 5, Appl 1
C 44	40	3.2	3366	1	US-08-469-802B-1	Sequence 1, Appl 1
C 45	40	3.2	3366	2	US-08-267-803B-1	Sequence 1, Appl 1

ALIGNMENTS

```

RESULT 1
US-09-253-691-3/C
Sequence 3, Application US/09253691
Patent No. 6124100
GENERAL INFORMATION:
APPLICANT: Dong Kyu JIN
TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
TITLE OF INVENTION: Using Trinucleotide Repeats Sequence
FILE REFERENCE: 1942/36
CURRENT APPLICATION NUMBER: US/09/253,691
CURRENT FILING DATE: 1999-02-22
EARLIER APPLICATION NUMBER: KR 98-6,278
EARLIER FILING DATE: 1996-02-26
NUMBER OF SEQ ID NOS: 3
SOFTWARE: WordPerfect 6.1/Windows
SEQ ID NO 3
LENGTH: 397
TYPE: DNA
ORGANISM: human
US-09-253-691-3

Query Match          4.2%; Score 51.8; DB 3: length 397;
Best Local Similarity 58.9%; Pred.No.2.5e-05;
Matches 89; Conservative 0; Mismatches 62; Indels 0; Gaps 0

QY      873 GCTGCTGCTGCTGCCGCCGCTGTTGTTGGCTGCACATGCTGCTGCTGTTGTTCT 932
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      200 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 141

QY      933 GCTGTACAGAAAAGAGCATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGA 992
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      140 GCTGTGCTCTCTTTTCTGCTGCTGCTGCAAAACATTCAAAAGTGAAGATATTTAAAAACAA 81

QY      993 CAAACAGAAAGTGAAGACAGAAAGTGAAGA 1023
        | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      80 AACTTAAGATAATTAACACATGAGAAAA 50

RESULT 2
US-08-531-927B-3/C
Sequence 3, Application US/0851927B
Patent No. 5840491
GENERAL INFORMATION:
APPLICANT: Kakizuka, Akira
TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph
Patent No. 5840491
TITLE OF INVENTION: Disease Gene and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

```

Db	296	CACGTATGCCAGTTACAAACACACATGACAGCTTCATCTCGAGTGTGATCATCATGATGT	355
Qy	355	GGACCCGAGTATGTTGGGGGAAACATCATATGAGCTGAGCTTCAGAACATGTCGGCTGATGATC	414
Db	356	GCACCCGATGACTCTCGGGATCCGTGCATATGAGCGCTCAGAACAGCATGGGTTTGATC	415
Qy	415	TGCTTACTTTCACCTGCTCCAAG	434
Db	416	TGCCTTCTCTCAGTGCAGG	435

RESULT 15	LOCUS	DEFINITION	663 bp	mRNA	linear	EST 21-JUL-2000
BE376590	BE376590	601226419p1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3584645 5'				
		mRNA sequence				

ACCESSION	BE370250
VERSION	BE376590.1
	GI:9321955

SOURCE	ORGANISM
house mouse.	Mus musculus.

REFERENCE 1 (bases 1 to 663)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1989)
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Gilbert Smith, Ph.D.
 CNA Library Preparation: Life Technologies, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MCC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: ILAB8743 row: f column: 06
 High quality sequence stop: 498.
 Location/Qualifiers
 1.663

```

/!AD_HOST="CHUB"
/!note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT      143 a      193 c      170 g      157 t
ORIGIN

```

```

Query Match      14.8%  Score 184;  DB 10;  Length 663;
Best local Similarity 73.4%  Pred. No. 3 4e-32;
Matches 235; Conservative 0; Mismatches 85; Indels 0; Gaps 0

```

QY AGGTTCTGGGTCGTGTATNTGAAGCATATGAAGGCCCCACAGTACCAATCCTCGAAGG 17
DB ACCTTCGGATCCACTATTATCAGATCATAGAAGGCTCTCAGATATTAACAATCCTTAAGGA 150

QY CTCCAGGCGTCCTTCAACTGCACCGTCTCCAGGSGCTGAAGCTCATCTGTGGGCTCT 234
DB CTACAGAGGCTCACTTCAACTGCACCGTGACTCACGGCTGGAAAGCTTCTCATGTGGACTCT 210

QY CAGTGACATGAGTGGGTAAAGGTGAGGCCCATGAGGCCATCATCCATGACCGCTT 294
DB TTAACCATAATGGTGGTGGCTGAGTCTTACCACCCAGAACCATCTACCAACAACCGCTT 270

QY CAACCTTCAGAGGTACGACAGAGGGGGAACCTTCACTCGAGATGATCATTCACAATGT 354

Db 271 CACCTTTGCCAGTTACACACGACACTGACGCTTCACTCTCGGAGTTGATCATTCATGATGT 330

QY 355 GGAGCCACGATGATTGGGGGACATCATGATGCACTCTCGAATCACTGCCCTGCATGATC 414

Db 331 GGAGGCCAGTAGTCGGGATCCGTCATGCACTGACACCTCGAAGACGCACTGGGTTCGGATC 390

QY 415 TGCCTTACCTTACCGTCCAG 434

Db 391 TGCCTTCCCTCATGATGCAAG 410

Search completed: April 29, 2003, 09:24:22
Job time : 2018.06 secs

```
Search completed: April 29, 2003, 09:24:22
Job time : 2018.06 secs
```

FEATURES	Location/Qualifiers
source	1. .1552

```

/note="data source:SPIR, source key:09NS15, evidence:ISS
homolog to IGFS PROTEIN (FRAGMENT)
putative"
/codon_start=1
/protein_id="BAB25436.1"
/db_xref="GI:12842010"
/db_xref="MGI:MGI:1919308"
/translation="MEGSRDPAVYVITIAQTASGSSVYIIEGNTAVYLDKSAHH
NCVTHGMMILMMTINOMVYVLSLTGGIITENRFTYASINSTDISTSLIHHQPS
DSSQVCSIONSHGSGSAFLSYQDSIGEEGPADEITMILILAAFAEISLLIIVITII
FCCCAKREKEESTYQNDIRKSAMRNKRADEPTKLSKEKYEVSDEAAQAOTAA
LPRKSEVSLPKRNSSLPYQELNKHDPGATPRVPSFDIASFOKRVNTYLL
1534...1539
polya_signal

```

[illegible][illegible]

REFERENCE
AUTHORS

TITLE
 JOURNAL
 COMMENT
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 www.washu.edu/nci/

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousestewartson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
this read has been verified (found to hit its original self in the
correct orientation)
MGI:427273
Seq primer: -40RP from Gibco
High quality sequence stop: 459
POLYA-No.

```

FEATURES
Source
Location/Qualifiers
1..486
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:693713"
/clone_1lb="Soares mouse NML"
/tissue_type="liver"
/lab_host="DH10B"
/note="Vector: pUT7AD-pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I oligo(dR) primer [5].
5'GTTACCACTCTGAGAGGAGGCGCCGCAATCTTTTATTTTATTTT 3';
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pUT73 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT
119 a 133 c 126 g 107 t 1 others

```

Query Match	14.9%	Score 184.6;	DB 9;	Length 486;
Best Local Similarity	73.4%;	Pred. 10, 2, 3e-32;		
Matches 225;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;
QY	115	AGGTCCTGGGCTCTGGTATATGATCAAGTCAATGAAAGGCCCCCAATGACCAACAGCTCCTGAAAGG	174	
Db	116	AGCTTCGGGATCCAGTATATGATGATCAATGAAAGGCTCTCAATATGTAACACATCTCCATGAAGGA	175	
QY	175	CTCCCAAGGCTCGCTCAACTGCAACCGCTCTCCACAGGGCTGGAAGCTCATCTGTGGCGCT	234	
Db	176	CTGAGAGGCTCACTTCAACTGCAACCGGATGCACTCCACGGCTGGAAAGCTTCTCATGTGGACTCT	235	
QY	235	CAGTGCATGTGTGTGTATAGCGTCAAGGGCCCATGAGAGCCCATCATCAACCAATACCGCTT	294	
Db	236	TACCAAAATGTGTGTGTGTGCTCAACCAACCCAAAGGACCATATATCAACCAACACCGCTT	295	
QY	295	CACCTCTCAGAGGTACGACCAAGGGCGGAACCTTCACTCGGAGATGATCATCCAAATGT	354	

MEDLINE 99279253

further

IMAGE Consortium (info@img.jml.gov) for further information
 MGI:940397
 Seq primer: -40RP from Gibco
 High quality sequence stop: 474.
 location/Qualifiers
 1. 611

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1513545"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="Female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pU73D-Pac (Pharmacia) with a modified
polylinker: 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pU73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonardo."

```

Query Match	15.0%	Score 185.6	DB 10	Length 611
Best Local Similarity	73.8%	Pred. No. 1.4e-32		
Matches 236	Conservative 0	Mismatches 84	Indels 0	Gaps 0
QY 115	AGGTTCTGGGCTGTGTAATGAAGTCATAGAAAGCCCCAGAAATGCAACAGTCTCTAAGGG	174		
Db 125	AGCTTCGGGATCCAGTTATCAGATCATAGAAAGGCTCTCAGATGTAAAGAGTCCCTTAAAGGA	184		
OY 175	CTCCGAGGCTCGCTTCACACTGCACCGCTCCCGAGGGTGGAAAGTCATCATGTGGCCTCT	234		
Db 185	CTCAGAGGCTCACTTCAACTGCACCGTCACTGACGCTGAGAGCTTCTCATSTGACACTCT	244		
OY 235	CAGTGACATGTGTGTCTTAAGGCGTCAGGCGCCATGGAGGCCATCATCACCAGATGACCGCTT	294		
Db 245	TAACCAAATGGTGTGTGTGATGAGTCTACACCAAGGAGGCCATCATCACCAGACACGCGTT	304		
OY 295	CACCTCTCAGAGGATCGACACGAGGGGGGAACTTCACCTCGGAGATGATCATCACAAATGT	354		
Db 305	CACCTATCCACAGTTTCAACAGCAGCATGACAGAGCTTCACTCTCGAGTGTGATCATCCAGATGT	364		
OY 355	GGAGCCGAGTGATGTTGGGGGAACATAGATGACAGCTCCGAGACAGTCCGCTCATGATGC	414		
Db 365	GCAGCCCACTAGTACTCGGGATCCGTGCATGACAGCTCGCAGAACACCAATGGSTTTGGATC	424		
OY 415	TGCTTACCTTACCGTCAAG	434		
Db 425	TGCTTCCCTCTCAGTCAAG	444		
RESULT 12				
LOCUS	Bg173684	931 bp	mRNA	linear
DEFINITION	602336623p1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:445973 5'			EST 06-FEB-2001
ACCESSION	Bg173684			
VERSION	Bg173684.1	GI:12680387		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	unpublished (1999)			
COMMENT	Contact: Robert Strusberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc.			

QY	971	AGGAAATCTGAAAAAGAGGAAGACAAAAGAACTGACACAGAAAGTGGAAATGAAAC	1030
QY	971	AGGAAATCTGAAAAAGAGGAAGACAAAAGAACTGACACAGAAAGTGGAAATGAAAC <td>1030</td>	1030
Db	377	AGGAAATCTGAAAAAGAGGAAGACAAAAGAACTGACACAGAAAGTGGAAATGAAAC	436
QY	1031	TCGGGCTACAAATTCAGATGAACAAAAGACCACAGAAACCGCTTCCTCCCTCCCAATTC	1090
Db	437	TCGGGCTACAAATTCAGATGAACAAAAGACCACAGAACCGCTTCCTCCCTCCCAATTC	496
QY	1091	TGTGAATCCAGTATCTCTGACAAAGAAACAGTAGCTGTGGCCCTCCCTACACAGCGGCT	1150
Db	497	TGTGAATCCAGTATCTCTGACAAAGAAACAGTAGCTGTGGCCCTCCCTACACAGCGGCT	556
QY	1151	GATCAACGTCACCCACAGCCAGCAAGTCATCCACAGGCTCTTTTAATCTGGCAAGTCT	1210
Db	557	GATCAACGTCACCCACAGCCAGCAAGTCATCCACAGGCTCTTTTAATCTGGCAAGTCT	616
QY	1211	GAGAGGTCAGTAATACAACTGTAGTATA	1239
Db	617	GAGAGGTCAGTAATACAACTGTAGTATA	645

```

FEATURES
SOURCE
LOCATION/Qualifiers
1..631
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E030016M12"
/clone_1lb="RIKEN full-length enriched, 0 day neonate
lung"
/tissue_type="Lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site 1: SalI; Site 2: BamHI: cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was

```


GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 23:25:02 ; Search time 2011.06 Seconds
(without alignments)
9985.963 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240
Sequence: 1 aggtgtgagtcacagcaaca.....gtatacaactgtatag 1240

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	308.2	24.9	784	12	BG663468
2	278.6	22.5	794	12	BG663468
3	265.8	21.4	725	12	BG740428
4	259	20.9	631	10	BG663870
5	246.2	19.9	527	10	BG032610
6	197.4	15.9	517	10	BE476432

7	190.8	15.4	622	17	AG069679
8	190.8	15.4	677	17	AG107877
9	190	15.3	524	10	AV601192
10	185.6	15.0	411	9	AA255274
11	185.6	15.0	611	10	AM990468
12	185.6	15.0	931	12	BG173684
13	185.6	15.0	1552	11	AK008060
14	184.6	14.9	468	9	A1425363
15	184	14.8	663	10	BE376590
16	182.4	14.7	650	10	BE625217
17	168.2	13.6	600	17	A2379623
18	154.8	12.3	556	12	BG206666
19	152.8	12.3	556	9	A1788300
20	151.2	12.2	530	9	A1790785
21	145.4	11.7	423	10	BB846577
22	143.8	11.6	291	17	AZ411779
23	128.8	10.4	471	12	BF040046
24	127	10.2	394	10	BB846133
25	124	10.0	685	17	AG142221
26	110	8.9	785	13	BI452873
27	102.8	8.3	889	13	BI454276
28	100	8.1	185	14	N47851
29	100	8.1	234	14	N93995
30	98.4	7.9	325	17	AQ403719
31	78.6	6.3	360	9	A1647601
32	78.6	6.3	374	9	A1648720
33	78.6	6.3	406	9	A1788973
34	77	6.2	644	10	BB513388
35	67.8	5.5	397	10	AM112084
36	63.2	5.1	282	10	BB564363
37	58.4	4.7	466	17	AQ224386
38	56	4.5	818	12	BG445021
39	52.6	4.2	422	9	A1756062
40	52.4	4.2	458	9	A1514929
41	50.6	4.1	639	17	AG113259
42	50	4.0	427	17	AO595624
43	50	4.0	447	10	AV431233
44	49.8	4.0	428	9	AA695064
45	49.6	4.0	925	17	CNS0381E

ALIGNMENTS

RESULT 1
BG663468
LOCUS
DEFINITION 602796941P1 NCI_GGAP_Mam4 Mus musculus cDNA clone IMAGE:4917803 5'

ACCESSION
BG663468
VERSION
BG663468.1 GI:14214006

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathli; Muridae; Murinae; Mus;

Ph.D.
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0830 row: b column: 12
High quality sequence stop: 788.
Location/Qualifiers


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5063 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA/DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 864..1349
FEATURE:
NAME/KEY: CDS
LOCATION: 1944..3370
FEATURE:
NAME/KEY: CDS
LOCATION: 3439..3736
US-08-185-432-1

```

```

Query Match          3.5%; Score 41; DB 1; Length 5063;
Best Local Similarity 66.3%; Pred. No. 0.076;
Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

```

QY 801 GCTGCTGCTGCTGCCGCGTGTGTTGTGCTGCAACTGCTGCGGTTGTTCT 860
      |||||
Db 2988 GCTGCTGCTGCTGCCGCGTGTGTTGTGCTGCAACTGCTGCGGTTGTTCT 2929
      |||||
QY 861 GCTGTAGAGAGAGAGATTCGTAT 889
      |||||
Db 2928 GCTGTGCTGCTGCCGATGCGATGAT 2900
      |||||

```

```

RESULT 15
US-08-479-913E-1/c
; Sequence 1, Application US/08479913E
; Patent No. 6416998
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Bert W.
; APPLICANT: Tsai, Ming-Jer
; APPLICANT: Ledebur, Harry C.
; APPLICANT: Kittle, Joseph D.
; TITLE OF INVENTION: MODIFIED STEROID HORMONES FOR GENE THERAPY AND METHODS FOR THEIR
; FILE REFERENCE: 212/133
; CURRENT APPLICATION NUMBER: US/08/479,913E
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 07/939,246
; PRIOR FILING DATE: 1992-09-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MicroSoft Word 97
; SEQ ID NO 1
; LENGTH: 6177
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Plasmid pGR0403R
US-08-479-913E-1

```

```

Query Match          3.5%; Score 40.8; DB 4; Length 6177;
Best Local Similarity 75.0%; Pred. No. 0.097;
Matches 51; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY 801 GCTGCTGCTGCTGCCGCGTGTGTTGTGCTGCAACTGCTGCGGTTGTTCT 860
      |||||
Db 4785 GCTGCTGCTGCTGCCGCGTGTGTTGTGCTGCAACTGCTGCGGTTGTTCT 4726
      |||||
QY 861 GCTGTAGA 868
      |||||
Db 4725 GCTGCACA 4718
      |||||

```

```

Search completed: April 29, 2003, 10:49:49
Job time : 81.8976 secs

```

```

      TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
      NUMBER OF SEQUENCES: 23
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: PENNIE & EDMONDS
      STREET: 1155 Avenue of the Americas
      City: New York
      STATE: New York
      COUNTRY: U.S.A.
      ZIP: 10036-2711

      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      OPERATING SYSTEM: IBM PC compatible
      SOFTWARE: Patentin Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/185,432
      FILING DATE: 21-JAN-1994
      CLASSIFICATION: 530
      ATTORNEY/AGENT INFORMATION:
      NAME: Mastrock, S. Leslie
      REGISTRATION NUMBER: 18,872
      REFERENCE/DOCKET NUMBER: 7326-006
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 790-9090
      TELEFAX: (212) 869-8864/9741
      TELEX: 66141 PENNIE
      INFORMATION FOR SEQ ID NO: 23:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 3771 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: unknown
      MOLECULE TYPE: cDNA
      US-08-185-432-23

      Query Match      3.5%; Score 41; DB 1; Length 3771;
      Best Local Similarity 66.3%; Pred. No. 0.064;
      Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

      QY 801 GCTGCTGCTGCTCCGCCGCTGTGTTGCTGCACACTGCTGCGCTGTGTTCT 860
          ||||||||| | | | | | | | | | | | | | | | | | | | | | | |
      Db 1897 GCTGCTGCTGCGCTTCTGCTGTGCACACTGCAGCTGTGCTGTTGTTGCTGT 1956
      QY 861 GCTGTGCAAGAAAGAGAGATTTCGTAAT 889
          ||||| | | | | | | | | | | | | | | | | | | | | | | | |
      Db 1957 GCTGTGCTGTCCGGATGCAGGTGAAT 1985

      RESULT 14
      US-08-185-432-1/c
      Sequence 1, Application US/08185432
      Patent No. 5750652
      GENERAL INFORMATION:
      APPLICANT: Artavanis-Tsakonas, Spyridon
      APPLICANT: Buseau, Isabelle
      APPLICANT: Diederich, Robert J.
      APPLICANT: Xu, Tian
      APPLICANT: Matsuno, Kenji
      TITLE OF INVENTION: DELTA PROTEINS, NUCLEIC ACIDS, AND
      TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
      NUMBER OF SEQUENCES: 23
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: PENNIE & EDMONDS
      STREET: 1155 Avenue of the Americas
      City: New York
      STATE: New York
      COUNTRY: U.S.A.
      ZIP: 10036-2711
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      OPERATING SYSTEM: IBM PC compatible
      SOFTWARE: Patentin Release #1.0, Version #1.30

```

PCT-US92-11107-12

Query Match

Best Local Similarity 64.9%; Score 42.6; DB 5; Length 3292;

Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 797 ATACGCTGCTGCTGCGCGCGTGTGTGTGCTGCAACTGCTGCGGTGTGTCT 856

DB 818 AGATGATATGTTGTTGTTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759

QY 857 TTCTGCTGTAGAGAAAAGAGATTGTGTATCAAT 893

DB 758 TGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722

RESULT 9

US-09-043-303-12/c

Sequence 12, Application US/09043303

Patent No. 6251589

GENERAL INFORMATION:

APPLICANT: SANPEI, Kazujiro

TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and

FILE REFERENCE: 0760-0241P

CURRENT APPLICATION NUMBER: US/09/043.303

EARLIER FILING DATE: 1998-05-18

EARLIER APPLICATION NUMBER: PCT/J996/01999

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 12

LENGTH: 78

TYPE: DNA

ORGANISM: Homo sapiens

US-09-043-303-12

Query Match

Best Local Similarity 78.1%; Score 41.6; DB 4; Length 78;

Matches 50; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 860

DB 67 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8

QY 861 GCTG 864

DB 7 GCTG 4

RESULT 10

US-09-043-303-7/c

Sequence 7, Application US/09043303

Patent No. 6251589

GENERAL INFORMATION:

APPLICANT: TSUJI, Shoji

TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and

FILE REFERENCE: 0760-0241P

CURRENT APPLICATION NUMBER: US/09/043.303

EARLIER FILING DATE: 1998-05-18

EARLIER APPLICATION NUMBER: PCT/J996/01999

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 7

LENGTH: 203

TYPE: DNA

ORGANISM: P-2093 plasmid

US-09-043-303-7

Query Match

3.6%; Score 41.6; DB 4; Length 203;

Best Local Similarity 78.1%; Pred. No. 0.0081;

Matches 50; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 801 GCTGCTGCTGCTGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 860

DB 67 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8

QY 861 GCTG 864

DB 7 GCTG 4

RESULT 11

US-08-145-617-5/c

Sequence 5, Application US/08145617

Patent No. 5766847

GENERAL INFORMATION:

APPLICANT: JACKIE, Herbert

TITLE OF INVENTION: PROCESS FOR ANALYZING LENGTH

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

STREET: 301 N. Washington Street, P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: United States of America

ZIP: 22046

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/145,617

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/6681,494

FILING DATE: 10-JUN-1991

APPLICATION NUMBER: DE P3834636.2

FILING DATE: 11-OCT-1988

ATTORNEY/AGENT INFORMATION:

NAME: Svensson, Leonard R.

REGISTRATION NUMBER: 30,330

REFERENCE/DOCKET NUMBER: 147-122PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-145-617-5

Query Match

Best Local Similarity 67.4%; Score 41.2; DB 1; Length 379;

Matches 58; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 779 CGAGCTGCTCTTACAAATACGCTGCTGCGCGCGTGTGTGTGTGTGTGTGTCT 838

DB 214 CCGAATCCAGGCGCTCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCT 155

QY 839 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864

DB 154 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129

```

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626
US-08-328-809-7

Query Match
Best Local Similarity 64.9%; Score 42.6; DB 1; Length 3292;
Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 797 ATACGCTGCTGCTGCTGCCGCCGCTGTTGTTGTGGCTGCACACTGCTGCTGCTTTGT 856
Db 818 AGATGATGATGTTGTTGTTGGAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGAC 759
OY 857 TTCTGCGTAGAAGAAAGAGATTGCTGTAAT 893
Db 758 TCTGCTGTTGCTGCACACTGCTGTTGTTGTTGTAATT 722

RESULT 8
PCT-US92-11107-12/c
Sequence 12, Application PC/PUS9211107
GENERAL INFORMATION:
APPLICANT: Donahue, Brian A.
APPLICANT: Toney, Jeffrey H.
APPLICANT: Bruhn, Suzanne L.
APPLICANT: Pil, Pieter M.
APPLICANT: Brown, Steven
APPLICANT: Kellett, Patti
APPLICANT: Essigmann, John M.
APPLICANT: Lippard, Stephen J.
TITLE OF INVENTION: DNA Structure Specific Recognition
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: 2 Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/11107
FILING DATE: 19921218
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539,906
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4787AAA
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3292 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Saccharomyces cerevisiae
IMMEDIATE SOURCE:
CLONE: lambda yPt
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1626

```



```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,927B
FILING DATE: 21-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP H6-251600
FILING DATE: 21-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: ATH95-01A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1776 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1115
US-08-531-927B-1

Query Match      3.7%; Score 42.8; DB 2; Length 1776;
Best Local Similarity 61.8%; Pred. No. 0.012;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 801 GCTGCGTCGTCGCCGCCCGTGTGGTGCTGCACAGTGGTGGCGGCTGTTGTTCC 860
    |||||.....|...|...|...|...|...|...|...|...|...|...|...|...|
Db 975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTT 916
    QY 861 GCTGTAGAAGAAAAAGAGAGATTCGTATTCAATTCCAAGAAATCTGAA 910
        |||||...|...|...|...|...|...|...|...|...|...|...|...|
        Db 915 GCTGCTGTTTTCGAAGTAGGCTTCTGCTCTTCCTCCGAGACTTTTGGAA 866
            |||||...|...|...|...|...|...|...|...|...|...|...|...|

RESULT 4
US-09-041-886-12/c
Sequence 12, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Redesen, Dale E.
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
```



```

Query Match: 4.4%; Score 51.8; DB 3; Length 397;
Best Local Similarity 58.9%; Pred. No. 1.2e-05;
Matches 89; Conservative 0; Mismatches 62; Indels 0; Gaps 0.

OY 801 GCTGCTGCTCTGCCGCCGTCGTTGTTGGCTGCAGACTGCTGCTGCCGTTGTTGTTCT 860
|||||
Db 200 GCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 141
|||||

OY 861 GCTGTGAGAAAGAAAAGAGATTTCGATTCATTTCAAAAGAAATTCGAAAAGACAGAGA 920
|||||
Db 140 GCTGTGCTGCTCTTTTGCTGCTGCTGTGCTGAACAATTCAGAAAGTAAAGTATATTTAAAAAACAA 81
|||||

OY 921 CAAACAAGAACTGAGACAGAAAGTGAAA 951
|||||
Db 80 AACTTAAGATTAATACACCATGAGAAA 50
|||||

RESULT 2
US-08-531-927B-3/C
; Sequence 3, Application US/0851927B
; Patent No. 5840491
; GENERAL INFORMATION:
; APPLICANT: Kakizuka, Akira
; TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph
; Patent No. 5840491
; TITLE OF INVENTION: Disease Gene and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;

```


XX Diagnosis of spinocerebellar ataxia type III (SCA III) syndrome
PT using techniques which ensure highly accurate diagnosis
XX

XX
XX
PA (SMSU) SAMSUNG MEDICAL CENT.
PA (JIND/) JIN D G.
XX
XX

PA (JIND/) JIN D G.
XX
PI
XX JIN DG;

SQ Sequence 7029 BP; 1426 A; 1440 C; 1945 G; 2218 T; 0 other;

Disclosure; Fig 10; 23pp; Korean.


```

QY 42 CCGTCTCGGCTCTGTAATGAGTCAATGAGGCCCCCAAAATGCAAGAGCTCTGAAG 101
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 108 CAGCTTCGGGATCCAGTATACATAGAGAGGCTCTCAGAAATGTAACGCTTAAGG 167
QY 102 GCTCCAGGCGTCCGTTCAACGTCACGCGTCCGAGGCTGGAAGCTCATCATGGGCTC 161
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 ACTGAGAGGCTCACTTCACTGACCGCTGACTCAGCGTGGAGCTTCTCATGTGGACTC 227
QY 162 TCAGTACATAGTGGTGTCTAAGCTCAGGCCCATGAGCCATCATCACCATGACCGCT 221
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 TTAACCAATAGTGGTGTGTAGTGTCTACCAACCCAGAGACCATCATCACCACACCGCT 267
QY 222 TCACCTCTCAGAGGTAGACACGAGGCGGGAATTCACCTGGAGATGATCCACAAAG 261
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 288 TCACCTATGCGAGTTACAAACGACACTGACAGCTTCATCTCGAGTGTATCATCATATG 347
QY 282 TGGAGCCAGTATGTCGGGGAAATGACATGACGCTCCAGCAAGACAGCGCTCATGATG 341
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 348 TGCAGCCAGTACTCGGGATCCGTGCAATGACGCTGCAAGAGCCATGAGGCTTGGAT 407
QY 342 CTGCTTACCTTACCGTCAAGATTATGGAGAGCTGTTCATTCCTCCAGTGTAACTGTAG 401
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 408 CTGCTTCTCTCAGTCAAGTCAATGGGGAACCTGAAACATTCCTAGCAACACCTATAG 467
QY 402 TCGCTGGAATGAACTTGTGAAGTACTGTCTACCTCAGACTGAGCTGGCTCCGG 461
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 468 TCACCTGAGGGGAAACCTGTAAATGTACTGTATGCGGCTGAGCTCCTCCCG 527
QY 462 ATATTCTCTGGAGCTCGTCTCTGTCAGCCATTCATTCATTTATTTTCCGAGC 521
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 528 ATATTCTCTGGAGCTGAGGTTCCCTTAAGCCATTCGATTAATTCCTTCTGAGC 567
QY 522 CCAAGCACTTCAAAATGACAGTACATCTGCTGACCCCAAGAGCAATGGAGCTT 561
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 588 CGGCAACTTATGAGGCTGTGAGTCTCTGAGACTCACACCACTGGGAGAGCGGAGCT 647
QY 582 TGACTTGGGGTACTGCTGAAGAGCTGAAAGCCCGCAAGTCTGCAACTGTAAATCTCA 641
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 648 TGACTTGTGGGAGAGCTGAAAGACTTGGAGCAGCAAGCTCTTAACGTCAACCTGA 707
QY 642 CTGTGATTCGGTCCCAAGACACTGAGAGTGTATTATATTCAGAGTATATCA 701
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 708 CTGTGATTCAGGCTCCACCTGACAGTATGGAAGA----- 744
QY 702 GTTTACGAGTATAGTTTTCATTCGCTACTTGGGCAAGTTGAGTGAAGTCAAG 761
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 745 -----AGGCCAGCACTGCGGCACTGCGGCATCATCTGCTGGAGTGGCCT 791
QY 762 GCACCATGCTTTCAGAGCCGACGTGACTTTCATTAATGCTGCTGCTGCCGCCGTC 821
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 792 TTTCTTGTCTCTTGAATCTGATCATGTTTGAATTAATATCTCTGTGTG----- 843
QY 822 GTTGTGTGCTCAACTGCTGCTGCGGTTGTTGTTCTCTGTAAGAAAGAGAT 881
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 844 -----TTGTGCTTCAGAGAGAGAAAGAGATCTACTTATATAAATGAAATAGAGAA 898
QY 882 TTGCTATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGTGAAGACAG 941
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 899 TCTGCAAACTGAGAGCA-----ACAAAGAGATCCGAGACAAAGTTAA 944
QY 942 AAGTGAATGAAACTCGGCTCAATTCAGATGAAACAAAGAGACAGACAGACCGCTT 1001
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 945 AAGTGAAGAAAGAAATTCAGGCTTACGATGAGGCAAGAGGCTGCAACAGCTGACAT 1004
QY 1002 CTCTCCCTCCCAATCTCTGAAATCCAGTATCTGAACAAAGAAAGACAGTGTGGCC 1061
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1005 CTCTCCCTCTTAATCTGCTGAAGTCAAGCTTCAGAAAGAGCAGACAGTA-----GCC 1058
QY 1062 CTCTCTACAGCGGGGCTGATCAAGCTCAACCCAGGCGAGAGTCAATCCACAGCTTCTT 1121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1059 TTTCTTATCAGGAATCAATTAACATCAGCCGCTGAGCAACTCATCCAGGCTTCTT 1118

```

```

QY 1122 TTAATCTGGCAGTCTCTGAGAGGTCAATATACAACTGTACTATATA 1167
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1119 TTGACATTCGCCAGTCTCTCAGAGGTCAAGATGCAAAATGTACTTATGTGA 1164
RESULT 7
AA136582
ID AA136582 standard; DNA; 401 BP.
XX
AC AA136582;
XX
DE 17-OCT-2001 (first entry)
XX
DE Probe #5268 used to measure gene expression in human placenta sample.
XX
DE Probe; microarray; human; placenta; antenatal diagnosis.
XX
DE genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00063.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0633365.
XX
PR 21-SEP-2000; 2000US-0233687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID NO 5268; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 other;
Query Match 16.8%; Score 196.8; DB 22; Length 401;
Best Local Similarity 96.6%; Pred. No. 1.6e-44;
Matches 201; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 27 AATATAGAGACCCACCCGCTTCTGGTCTGTAATGAGTCAATGAGGCCCCCAAAATG 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 AATATGGCTACTCTCCAGTGTCTGTGTAATGAGTCAATGAGGCCCCCAAAATG 253
QY 87 CAAGAGTCTGAGAGGCTCCAGAGCTGCTTCAACTGACCGCTCCAGGCTGGAAGC 146
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 254 CAAGAGTCTGAGAGGCTCCAGAGCTGCTTCAACTGACCGCTCCAGGCTGGAAGC 313
QY 147 TCATATGTGGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 206
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 314 TCATATGTGGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
QY 207 TCACCAATGACCGCTTCACTCTCAGAG 234
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 374 TCACCAATGACCGCTTCACTCTCAGAG 401

```


Db	241	GGGAACCTCACCTGGAGATGATCATCCAAATGTGGAGCCCGAGTATTCGGGGAACATC	300
QY	308	AGATGACACCTCCGCAAGACAGCGCCCTGCATGATGATTCGCTTACCTTACCGTCCAAAGTTAG	367
Db	301	AGATGACACCTCCGCAAGACAGCGCCCTGCATGATGATTCGCTTACCTTACCGTCCAAAGTTAG	360
QY	368	GGAGAGCTGTTCATTCCCACTGCTTAACTCTGTAGTGCCTGAGAAATGAACTCTTGAGTT	427
Db	361	GGAGAGCTGTTCATTCCCACTGCTTAACTCTGTAGTGCCTGAGAAATGAACTCTTGAGTT	420
QY	428	ACTTGTCTACCTTCACTGAGCTGGCTCCCGGATATTTCTGGAGCTCGTCTCTG	487
Db	421	ACTTGTCTACCTTCACTGAGCTGGCTCCCGGATATTTCTGGAGCTCGTCTCTG	480
QY	488	GTCAGCACTTCAAGCTATTTATTTGTTCCGAGGCCAGCAGCACTTAAAGTGAAGGAC	547
Db	481	GTCAGCACTTCAAGCTATTTATTTGTTCCGAGGCCAGCAGCACTTAAAGTGAAGGAC	540
QY	548	ATTCCTGGCTCTGACCCCAAGAGCAATGGACATTGACTTTGCTGGTACTGGAAGAC	607
Db	541	ATTCCTGGCTCTGACCCCAAGAGCAATGGACATTGACTTTGCTGGTACTGGAAGAC	600
QY	608	CTGAAGGCCCCCAAGTGCAGCACTGATAATCTACTGTGATTCGGTCCCAAGCACT	667
Db	601	CTGAAGGCCCCCAAGTGCAGCACTGATAATCTACTGTGATTCGGTCCCAAGCACT	660
QY	668	GGAGGTGATTAAATATTCAGAGTATTTATCAAGTTTACAGATTAGATTTTTCATTG	727
Db	661	GGAGGTGATTAAATATTCAGAGTATTTATCAAGTTTACAGATTAGATTTTTCATTG	720
QY	728	CCTACTGGGGGCAAGTTGAGACTGGACCTGACAGGACCACTGTTTGACGGCAGCTG	787
Db	721	CCTACTGGGGGCAAGTTGAGACTGGACCTGACAGGACCACTGTTTGACGGCAGCTG	780
QY	788	ACTCTTACAAATACGCTCTGCTGCTGCCGCGTCTGTTGTTGGCTGCAACTGCTGTC	847
Db	781	ACTCTTACAAATACGCTCTGCTGCTGCCGCGTCTGTTGTTGGCTGCAACTGCTGTC	840
QY	848	CGTTGTTGTTCTCTCTGTAGAAAGAAAGGATTTCGATTTCAATTTCAAAAAGAACT	907
Db	841	CGTTGTTGTTCTCTCTGTAGAAAGAAAGGATTTCGATTTCAATTTCAAAAAGAACT	878
QY	908	GAAGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGCTAC	967
Db	879	GAAGAAAGAGAGACAAACAAAGAACTGAGACAGAAAGTGAATGAAACTCCGGCTAC	938
QY	968	AATTCAGATGAACAAAGACCAAGACACCGCTTCTCTCCCTCCCAATCTGTGAATCC	1027
Db	939	AATTCAGATGAACAAAGACCAAGACACCGCTTCTCTCCCTCCCAATCTGTGAATCC	998
QY	1028	AGTATCTCTGAACAAAGAAACAGTACGTGTGGCCCTCTCACCAGGGGGCTATCAAGT	1087
Db	999	AGTATCTCTGAACAAAGAAACAGTACGTGTGGCCCTCTCACCAGGGGGCTATCAAGT	1058
QY	1088	CCACCCAGGCGCAGCAATCTCAATCCACAGGCTTCTTTAACTGTGCGAGTCTGAGAAGTC	1147
Db	1059	CCACCCAGGCGCAGCAATCTCAATCCACAGGCTTCTTTAACTGTGCGAGTCTGAGAAGTC	1118
QY	1148	AGTAATACACTGTAGATA 1167	
Db	1119	AGTAATACACTGTAGATA 1138	
RESULT 5			
AAS92356			
ID	AAS92356 standard; cDNA; 1392 BP.		
XX	AAS92356;		
XX	13-FEB-2002 (first entry)		
XX	DNA encoding novel human diagnostic protein #28160.		

KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KV	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	W0200175067-A2.
XX	
PD	11-Oct-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YF;
XX	
DR	WPI: 2001-639362/73.
DR	P-PSDB: ABG28169.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -

Claim 1; SEQ ID No 28160; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at <http://wipo.int/pub/pct/> sequences.

SQ Sequence 1392 BP; 338 A; 361 C; 341 G; 352 T; 0 other;

Query Match	Score	DB	Length
66.68;	777.6;	23;	1392;

Matches 912; Conservative 0; Mismatches 24; Indels 110; Gaps 1

QY 152 ATGTGGCTCTCAGTGACATGGTGGTGAAGCGTCAGGCCATGGAGCCCATCATCACC 211

Db 1 ATGTGGCTCTCAGTGACATGGTGGTGAAGCGTCAGGCCCATGGAGCCCATCATCACC 60

QY 212 AATGACCGCTCACCCTCAGAGGTACGACCAGGGCGGAACCTCACCCTCGAGATGATC 271

Db 61 AATGACCGCTTCACTCTCAGAGGTACGACCAGGGCGGAACTTCACCTCGGAGATGATC 120

272 ATCCACAATGTGGAGCCAGTGATTGGGGAACATCAGATGCAGCCTCCAGAACAGTCGC 331

Db 121 ATCCACAATGTGGAGCCAGTGATTCTGGGGACATCAGATGCAGCCTCCAGAACAGTCC 180

332 CTGCATGGATCTGCTTACCTTACCGTCCAGTTATGGGAGAGCTGTTCAATCCCACTGT 391

Db 181 CTGCATGGATCTGCTTACCTTACCGTCCAAGTTATGGGAGAGCTGTTCATTCCCACTGTT 240

392 ATCTGTAGTCGCTGAGATGACCTTGTAAGTTACTTGTCTACCCCTCACACTGGACC 451

QY 634 AATCTCACTGATTCGATGTCCTCCCAAGACACTGAGTGTATTAATATTCAGGTGT 693
 |||||||
 Db 706 AATCTCACTGATTCGATGTCCTCCCAAGACACTGAGTGTATTAATATTCAGGTGT 765
 QY 694 ATATCAAGTTTACCGATTAGGTTTTCATTGCTTACTTGGGCAAGTGGACTTGG 753
 |||||||
 Db 766 ATATCAAGTTTACCGATTAGGTTTTCATTGCTTACTTGGGCAAGTGGACTTGG 825
 QY 754 ACTAGCAGGACCATGCTTCTGAGCCGAGCTGCTACTTACATATAGCTGCTGCTG 813
 |||||||
 Db 826 ACTAGCAGGACCATGCTTCTGAGCCGAGCTGCTACTTACATATAGCTGCTGCTG 885
 QY 814 CCGCCGCTGCTGTTGGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873
 |||||||
 Db 886 CCGCCGCTGCTGTTGGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
 QY 874 AAGAGATTTCGTATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAC 933
 |||||||
 Db 946 AAGAGATTTCGTATTCATTTCAAAAGAAATCTGAAAAGAGAGACAAACAAAGAAC 1005
 QY 934 TGACACAGAAAGTGAATAAGAAATCCGCTACATTCAGATGAACAAAGACACAGA 993
 |||||||
 Db 1006 TGACACAGAAAGTGAATAAGAAATCCGCTACATTCAGATGAACAAAGACACAGA 1065
 QY 994 CACGCTCTCTCCCTCCCAATCTCTGATTCAGATTCAGATTCAGATTCAGATTCAG 1053
 |||||||
 Db 1066 AACGCTCTCTCCCTCCCAATCTCTGATTCAGATTCAGATTCAGATTCAGATTCAG 1125
 QY 1054 CTGTGGCCCTCTCCACAGCGGGCTGATTCACAGTCCACCCAGGCGACAGTCATCCACA 1113
 |||||||
 Db 1126 CTGTGGCCCTCTCCACAGCGGGCTGATTCACAGTCCACCCAGGCGACAGTCATCCACA 1185
 QY 1114 GCGTTCTTTTATTCGCGCACTCTCTGAGAGGTAGTATACACTGATGATG 1168
 |||||||
 Db 1186 GCGTTCTTTTATTCGCGCACTCTCTGAGAGGTAGTATACACTGATGATG 1240
 RESULT 4
 ABK13031
 ID ABK13031 standard; cDNA: 1139 BP.
 AC ABK13031;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding human B7-1-like protein, B7-L_{h4}.
 XX
 KW Human; B7-1-like protein; B7-L; antifertility; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KW antidiabetic; haemostatic; antithyroid; antitumor; antiallergic;
 KW antiautismal; nephrotropic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1134
 FT /*tag= a
 FT /product= "B7-1-like protein, B7-L_{h4}"
 XX
 PN MO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PE 28-JUN-2001; 2001MO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX

PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Samiento UM, Schultz HJ, Chute HT;
 XX
 DR WPI; 2002-130881/17.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 1; Fig 4; 135pp; English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allosensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L_{h4}.
 CC
 XX
 SO Sequence 1139 BP; 290 A; 300 C; 283 G; 266 T; 0 other;
 Query Match 94.4%; Score 1102.8; DB 24; Length 1139;
 Best Local Similarity 97.9%; Pred. No. 1.9e-296;
 Matches 1136; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
 QY 8 ATGTGTGAGAGGAGCCATGGAATAAGAGACCCAGCGTTCTGGCTGTGTAATGAGTGC 67
 |||||||
 Db 1 ATGTGTGAGAGGAGCCATGGAATAAGAGACCCAGCGTTCTGGCTGTGTAATGAGTGC 60
 QY 68 ATAGAGAGGCCCCCAAAATGCAAGAGTCTGAAAGGCTCCAGGCTGCTCACTGACACC 127
 |||||||
 Db 61 ATAGAGAGGCCCCCAAAATGCAAGAGTCTGAAAGGCTCCAGGCTGCTCACTGACACC 120
 QY 128 GTCTCCAGGCGTGAAGGTCATCATGCTGAGGCTCTGAGTATGAGTGTGTAAGCGGC 187
 |||||||
 Db 121 GTCTCCAGGCGTGAAGGTCATCATGCTGAGGCTCTGAGTATGAGTGTGTAAGCGGC 180
 QY 188 AGGCCCATGAGGCCATCATCAATGACCGCTTCACTCTCAGAGTACAGCAGGCGC 247
 |||||||
 Db 181 AGGCCCATGAGGCCATCATCAATGACCGCTTCACTCTCAGAGTACAGCAGGCGC 240
 QY 248 GGGAACTTCACTCGAGATGATCATCAATGTGAGAGCCAGTATTCGGGGAACATC 307
 |||||||

seminal vesicle hyperplasia in transgenic mice overexpressing B7-L polypeptide. Hence modulators of (1) are useful for the treatment of cancer including seminal vesicle cancer, lung, brain, breast, ovarian, testicular cancer and cancers of hematopoietic system. B7-L polypeptide pathway can be manipulated to regulate cytotoxic T-lymphocyte response in allograft transplantation, graft versus host disease, T-cell dependent B-cell mediated diseases and autoimmune diseases. B7-L molecules are useful for alleviating the symptoms associated with diseases involving chronic immune cell dysfunction or to treat autoimmune diseases such as systemic lupus erythematosus, rheumatoid arthritis, multiple sclerosis, diabetes, immune thrombocytopenic purpura and psoriasis, chronic inflammatory disease such as inflammatory bowel disease (Crohn's disease and ulcerative colitis), Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They are also useful as immunosuppressive agents for bone marrow and organ transplantation or to prolong graft survival. B7-L molecules are also useful for diagnosis and treatment of diseases involving abnormal cell proliferation, including arteriosclerosis and vascular restenosis. Antagonists of B7-L polypeptides are useful for alleviation of toxic shock syndrome or allo sensitisation due to blood transfusions, and for treatment of allergy, asthma and hypersensitivity reactions, nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and pemphigoid), endocrinopathies (Grave's disease), various pneumopathies (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and lymphoproliferative disorders such as multiple myeloma. The present sequence represents the coding sequence of human B7-L_{h3}.

Sequence 1240 BP; 318 A; 319 C; 305 G; 298 T; 0 other;

Query Match	95.8%	Score 1119;	DB 24;	Length 1240;
Best local Similarity	99.1%;	Pred. No. 6.1e-301;		
Matches 1125;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0
34 AGACCCACC	CGGCTTGGGCTTGGTATGAAGTCATAGAGGCCCCCAAAATGCAAGCT	93		
Db 106 AGAAGCTG	TAGATTCTGGGGCTGTGATGAAGTCATAGAGGCCCCCAAGATGCAACAGT	165		
94 CCGAAGAGG	GCTCCAGAGCTCGCTTCAACAGCCACCGCTCCAGAGGCTGGAAGTCTATCAT	153		
Db 166 CCGAAGAGG	GCTCCAGAGGCTCGCTTCAACAGCTACCGCTCCAGAGGCTGGAAGTCTATCAT	225		
9Y 154 GTGGGCTC	AGTACATGATGGTGTGCTTAAGCTCAGAGCCCATGAGACCCATCATCCAA	213		
Db 226 GTGGGCTC	CAGACATGATGGTGTGCTTAAGCGTCAAGGCCCATGAGACCCATCATCCAA	285		
9Y 214 TGACCGCTT	CACCTCTCAAGAGTACACAGGCGGGAACTTCACTCGGAATGATCAT	273		
Db 286 TGACCGCTT	CACCTCTCAAGAGTACACAGGCGGGAACTTCACTCGGAATGATCAT	345		
9Y 274 CCACATGTG	AGAGCCAGTATGGGGGAAATCAAGATCAGACCTCCAAAGATCGCT	333		
Db 346 CCACATGTG	AGAGCCAGTATGGGGGAAATCAAGATCAGACCTCCAAAGATCGCT	405		
9Y 334 GCATGAT	TGCTTACCTTACCGTCAAGTATVGGAGAGAGCTTTCATCCAGTGTAA	393		
Db 406 GCATGAT	TGCTTACCTTACCGTCAAGTATVGGAGAGAGCTTTCATCCAGTGTAA	465		
9Y 394 TCTTGAG	TCGTGAGAAATGAACCTTGGAAATTAATCTGTCTACCTCAGATGACCTG	453		
Db 466 TCTTGAG	TCGTGAGAAATGAACCTTGGAAATTAATCTGTCTACCTCAGATGACCTG	525		
9Y 454 GCTCCG	GAATATTTCTCTGGAGACTTGATCTCTCTGCTCAGACCTCAACGATTAATTTGT	513		
Db 526 GCTCCG	GAATATTTCTCTGGAGACTTGATCTCTCTGCTCAGACCTCAACGATTAATTTGT	585		
9Y 514 TCCGGAG	CCGACGACCTTCAAATGTCAGTACATCTGGCTGTGACCCACAGAGCAA	573		
Db 586 TCCGGAG	CCGACGACCTTCAAATGTCAGTACATCTGGCTGTGACCCACAGAGCAA	645		
9Y 574 TGGGACTT	TGACTTGCCTGGGCTACTGTGAAGAGCCTGAAGGCCCGCAAGTGTGCACTGT	633		
Db 646 TGGGACTT	TGACTTGCCTGGGCTACTGTGAAGAGCCTGAAGGCCCGCAAGTGTGCACTGT	705		

DT 23-APR-2002 (first entry)
 XX DNA encoding human B7-1-like protein, B7-L_h1.
 DE
 XX Human; B7-1-like protein; B7-L; antiinfectivity; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antirheumatic;
 KW antinflammatory; dermatological; antiproliferative; neuroprotective;
 KW antidiabetic; haemostatic; antihypertoid; antileuk; antiallergic;
 KW antitubercular; nephrotoxic; antibacterial; vitruce; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 27.1175
 FT CDS /tag= a
 FT /product= "B7-1-like protein, B7-L_h1"
 FT
 PN WO200200710-A2.
 XX 03-JAN-2002.
 XX 28-JUN-2001; 2001WO-US20719.
 XX 28-JUN-2000; 2000US-214512P.
 XX 28-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 PA
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX WPI: 2002-130881/17.
 DR P-SDB: AA075540.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 XX Claim 1; Fig 1; 135pp; English.
 PS
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility) and
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or alloimmunisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and

CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia, and
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L_h1.
 XX
 SO Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other;
 Query Match 96.1%; Score 1122.4; DB 24; Length 1175;
 Best Local Similarity 99.5%; Pred. No. 6.8e-302;
 Matches 1126; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 37 CCCACCCGCTTCGGTCTGTAATGAAGTCAAGAGCCCAAAATGCAAGAGTCT 96
 44 CCTCCACGCTTCGGTCTGTAATGAAGTCAAGAGCCCAAAATGCAAGAGTCT 103
 97 GAAGGCTCCAGGCTGCTTCAACTGACCGTCTCCAGGCTGGAAGTCAATG 156
 104 GAAGGCTCCAGGCTGCTTCAACTGACCGTCTCCAGGCTGGAAGTCAATG 163
 157 GCGTCTGAGTCAATGGTGTCTAAGCGTCAAGCCCATGAGCCATCATCACA 216
 164 GCGTCTGAGTCAATGGTGTCTAAGCGTCAAGCCCATGAGCCATCATCACA 223
 217 CCGCTTCACTCTCAGAGTACGACGAGCGGGAACCTTCACTTCGAGATCATCA 276
 224 CCGCTTCACTCTCAGAGTACGACGAGCGGGAACCTTCACTTCGAGATCATCA 283
 277 CAATGTGAGCCGAGTATGGGGGACATGATGAGTCAAGTCAAGTCAAGTCA 336
 284 CAATGTGAGCCGAGTATGGGGGACATGATGAGTCAAGTCAAGTCAAGTCA 343
 337 TGGATCTGCTTACCTTACCGTCAAGTATGAGAGTCTTCACTTCGAGATCAT 396
 344 TGGATCTGCTTACCTTACCGTCAAGTATGAGAGTCTTCACTTCGAGATCAT 403
 397 TGTAGTCTGAGTATGAGTCAAGTATGAGAGTCTTCACTTCGAGATCAT 456
 404 TGTAGTCTGAGTATGAGTCAAGTATGAGAGTCTTCACTTCGAGATCAT 463
 457 CCGGATATTTCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
 464 CCGGATATTTCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523
 517 GGAGCCAGCAGCTTCAAAATGACATGAGTCAAGTCTGAGTCAAGTCAAG 576
 524 GGAGCCAGCAGCTTCAAAATGACATGAGTCAAGTCTGAGTCAAGTCAAG 583
 577 GACTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
 584 GACTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 643
 637 TCTCACTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 696
 644 TCTCACTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703
 697 ATCACTTACGAGTATGATTTTCACTTCTGCTGCTGCTGCTGCTGCTGCT 756
 704 ATCACTTACGAGTATGATTTTCACTTCTGCTGCTGCTGCTGCTGCTGCT 763
 757 AGAGGACACAGCTTCTGAGCGGCGAGTCTGCTTCAATATGCTGCTGCTG 816
 764 AGAGGACACAGCTTCTGAGCGGCGAGTCTGCTTCAATATGCTGCTGCTG 823
 817 CCGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
 824 CCGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
 877 AGGATTTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAAG 936
 884 AGGATTTTCTGATTTCAATTTCAAAAGAAATCTGAAAAGAGAGACAAAG 943
 937 GACAGAAAGAGAAATGAAAGTCCGGCTACAAATTCAGATGAAAGAAAGAC 996

PR 28-JUN-2000; 2000US-214512P.
PR 28-NOV-2000; 2000US-0729264.
XX
PA (AMGE-) AMGEN INC.
XX
XX
PI Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
DR WPI: 2002-130881/17.
XX P-PADB: AAO75541.
DR
XX New B7-like polypeptides, polynucleotides and their modulators, useful
PT for diagnosing, preventing and treating reproductive, immune and
PT proliferative disorders, e.g. cancer and arteriosclerosis -
XX
XX Claim 1; Fig 2; 135pp; English.

Sequence 1168 BP; 299 A; 302 C; 288 G; 279 T; 0 other;

Query Match	100.0%;	Score 1168;	DB 24;	Length 1168;
Best Local Similarity	100.0%;	Pred. No. 1.1e-314;		
Matches 1168; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	AGTATCATGTTGGCAGAGACCATGAAAAATAGAGACCACCCGGTCTGGGTCTGGTAA	60
Db	1	AGTATCATGTTGGCAGAGACCATGAAAAATAGAGACCACCCGGTCTGGGTCTGGTAA	60
QY	61	TGAAGTCATAGAAAGCCCCCAAAATGCAAGTCTCAAGGGCTCCAGGCTGGCTTCAA	120
Db	61	TGAAGTCATAGAAAGCCCCCAAAATGCAAGTCTCAAGGGCTCCAGGCTGGCTTCAA	120
QY	121	CTGACACGTCCTCCAGGGCTGGAAAGTCATCATCTGGGCTCTCAGTACATGGTGGTCT	180
Db	121	CTGACACGTCCTCCAGGGCTGGAAAGTCATCATCTGGGCTCTCAGTACATGGTGGTCT	180
QY	181	AAGCGTAGAGGCCCATGGAGGCCATCATACCAATGACCGCTTCACTCTCAGAGGTACGA	240

Db	181	AAGGCTCAGAGCCCATGGAGCCCATCATCAACAAATGACCGCTTCACCTCTCAGAGTACGA	240
Qy	241	CCAGAGGGGGAACCTTACACCTCGGAGATGATCATCCAAATGAGAGCCACGATGATCGG	300
Db	241	CCAGAGGGGGAACCTTACACCTCGGAGATGATCATCCAAATGAGAGCCACGATGATCGG	300
Qy	301	GAAATCAGATGAGAGCCTCCAGAAACAGTCGCTGCATGATTCGTTACCTTACCGTCA	360
Db	301	GAAATCAGATGAGAGCCTCCAGAAACAGTCGCTGCATGATTCGTTACCTTACCGTCA	360
Qy	361	AGTTATGGGAGAGCTGTTCAATTCGCCAGTGTAACTTGTAGTCGTGAAGATGAACCTTG	420
Db	361	AGTTATGGGAGAGCTGTTCAATTCGCCAGTGTAACTTGTAGTCGTGAAGATGAACCTTG	420
Qy	421	TGAAGTACTAGTTCCTACCCACACAGGACCTGGCTCCCGGATTTCTCTGGAGACTCG	480
Db	421	TGAAGTACTAGTTCCTACCCACACAGGACCTGGCTCCCGGATTTCTCTGGAGACTCG	480
Qy	481	TCCTCTGGTAGCCATTCGAAGTATTAATTTTGTCCGAGCCACAGCACTTCAAAATGC	540
Db	481	TCCTCTGGTAGCCATTCGAAGTATTAATTTTGTCCGAGCCACAGCACTTCAAAATGC	540
Qy	541	AGTAGACATCCTGGCTCTGACCCCAAGAGCAATGGGAATTGACTTGCATGGCTACCTG	600
Db	541	AGTAGAGATCCTGGCTCTGACCCCAAGAGCAATGGGAATTGACTTGCATGGCTACCTG	600
Qy	601	GAAAGAGCTGAAAGCCCGGCAAGTCTCAACTGTAAATCTCACTGTGATTCGGTGTCCCA	660
Db	601	GAAAGAGCTGAAAGCCCGGCAAGTCTCAACTGTAAATCTCACTGTGATTCGGTGTCCCA	660
Qy	661	AGACACTGAGAGTGTATTAATTAATTCAGGTATTAATCAAGTTTACGAGATTTAGTTT	720
Db	661	AGACACTGAGAGTGTATTAATTAATTCAGGTATTAATCAAGTTTACGAGATTTAGTTT	720
Qy	721	TTCAATTCGCTACTGGGGGCAAAAGTGAAGTGGACTGACTAGAGGACACATGCTGTACCC	780
Db	721	TTCAATTCGCTACTGGGGGCAAAAGTGGACTGACTAGAGGACACATGCTGTACCC	780
Qy	781	GAGCTGTACTCTTACAAATACGCTGTGCTGCGCGCGTGTGTTGTGGCTGCACTG	840
Db	781	GAGCTGTACTCTTACAAATACGCTGTGCTGCGCGCGTGTGTTGTGGCTGCACTG	840
Qy	841	CTGTCGCGCGTGTGTTGTGCTGCTGCTGGAAGAAAGAAAGAGATTTGATTAATTCACAA	900
Db	841	CTGCTCGCGCTGTGTTGTGCTGCTGCTGGAAGAAAGAAAGAGATTTGATTAATTCACAA	900
Qy	901	GAAATCTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGAATGAAAACTC	960
Db	901	GAAATCTGAAAAAGAGAAGCAAAACAAAGAACTGAGACAGAAAGTGAATGAAAACTC	960
Qy	961	CGGCTACAAATTCAGATTAACAAAAAGCACAGACCGCTTCTCTCCGCCCAAAATCTG	1020
Db	961	CGGCTACAAATTCAGATTAACAAAAAGCACAGACCGCTTCTCTCCGCCCAAAATCTG	1020
Qy	1021	TGAATTCAGATGATCTGTAACAAAGAAAGAGTACTGTGGCCCTCCCTCAACACAGCGGCTGA	1080
Db	1021	TGAATTCAGATGATCTGTAACAAAGAAAGAGTACTGTGGCCCTCCCTCAACACAGCGGCTGA	1080
Qy	1081	TCAACTGCACCCACAGGCGCAGCAAGTCATCCACAGGCTTTTAAATCTGCGCATGCTCGA	1140
Db	1081	TCAACTGCACCCACAGGCGCAGCAAGTCATCCACAGGCTTTTAAATCTGCGCATGCTCGA	1140
Qy	1141	GAAAGTCAATGATATCACTCACTGTAGATATG 1168	
Db	1141	GAAAGTCAATGATATCACTCACTGTAGATATG 1168	
RESULT 2			
ID	ABK13028	standard; cdNA; 1175 BP.	
NC	ABK13028;		
XX			

```
RESULT 2
ABK13028
ID ABK13028 standard; CDNA; 1175 BP
XX
AC ABK13028;
```


C	10	53.8	4.6	277	21	ABK10239	Trinucleotide repeat
C	11	52.4	4.5	7029	23	ABL29756	Drosophila melanog
C	12	51.8	4.4	389	21	ABK10248	Tri-nucleotide rep
C	13	51.8	4.4	397	20	MAX88991	Spinoerebellar at
C	14	51.8	4.4	403	21	ABK10240	Spinoerebellar at
C	15	49.4	4.2	46954	23	ABL16830	Rat androgen rece
C	16	48.4	4.1	3217	10	AAN91578	Rat androgen rece
C	17	48.4	4.1	3217	12	AAO12002	Full-length rat an
C	18	48.4	4.1	4180	10	AAN91373	Rat androgen rece
C	19	47.8	4.1	381	23	ABV54466	Human prostate exp
C	20	47.6	4.1	273	20	MAX84442	Mouse brain CNG-1
C	21	47.4	4.1	3894	23	ABL03353	Drosophila melanog
C	22	47.4	4.1	22341	23	ABL03352	Drosophila melanog
C	23	46.6	4.0	2333	22	AAF75339	Human TGF-beta rec
C	24	46.4	4.0	1983	23	ABL24483	Drosophila melanog
C	25	46.4	4.0	4044	23	ABL24482	Drosophila melanog
C	26	45.8	3.9	2682	23	ABL16097	Drosophila melanog
C	27	45.8	3.9	8821	23	ABL16096	Drosophila melanog
C	28	45.6	3.9	1509	23	ABL24571	Drosophila melanog
C	29	45.6	3.9	3884	23	ABL24570	Drosophila melanog
C	30	45.4	3.9	3042	23	ABL28445	Drosophila melanog
C	31	45.4	3.9	3135	23	ABL28653	Drosophila melanog
C	32	45.4	3.9	5215	23	ABL28444	Drosophila melanog
C	33	45.4	3.9	5447	23	ABL28652	Drosophila melanog
C	34	45.2	3.9	5397	23	ABL29757	Drosophila melanog
C	35	45	3.9	462	23	ABL28941	Drosophila melanog
C	36	45	3.9	2115	23	ABL26940	Drosophila melanog
C	37	45	3.9	2260	23	ABL03191	Drosophila melanog
C	38	45	3.9	3210	23	ABL08977	Drosophila melanog
C	39	45	3.9	3624	23	ABL03190	Drosophila melanog
C	40	45	3.9	5975	23	ABL08976	Drosophila melanog
C	41	45	3.9	6541	23	ABL04172	Drosophila melanog
C	42	44.8	3.8	254	21	AAZ45758	Polyomphic region
C	43	44.8	3.8	1310	20	AAV08558	Transgene for epit
C	44	44.8	3.8	2614	24	ABR36100	CDNA sequence #491
C	45	44.8	3.8	3263	20	AAV08560	Transgene for epit

ALIGNMENTS

RESULT 1	
ID	ABK13029
ABK13029	standard; cdna; 1168 BP.
XX	ABK13029;
XX	AC
XX	DT
XX	23-APR-2002 (first entry)
DE	DNA encoding human B7-like protein, B7-L _{h2} .
XX	
KW	Human; B7-like protein; B7-L; antiinferility; gynaecological;
KW	antitumour; cytoskeletal; immunosuppressive; antiarthritis; antineumatic;
KW	antiinflammatory; dermatologic; antiproliferative; neuroprotective;
KW	antidiabetic; haemostatic; antihypertensive; antiallergic;
KW	antisthmatic; nephrotoxic; antibacterial; virucide; tumour; cancer;
KW	reproductive disorder; graft versus host disease; autoimmune disease;
KW	toxic shock syndrome; allergy; nephropathy; skin disorder;
KW	endocrinopathy; lymphoproliferative disorder; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	8..1168
FT	/tag= a
XX	/product= "B7-like protein, B7-L _{h2} "
XX	
XX	WO200200710-A2.
XX	
PD	03-JAN-2002.
XX	
PF	28-JUN-2001. 2001WO-US20719.
XX	

[illegible]

Meisterernst, Michael
Polites, Greg
TITLE OF INVENTION: PURIFICATION OF HIGHER ORDER TRANSCRIPTION
COMPLEXES FROM TRANSGENIC
NON-HUMAN ANIMALS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: HELLER, EHRMAN, WHITE & MCAULIFFE
STREET: 1666 K Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/849,243
FILING DATE: 07-May-2001
ATTORNEY/AGENT INFORMATION:
NAME: Granados, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 38005-0148
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)912-2000
TELEFAX: (202)912-2020
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: exon
LOCATION: 1..3263
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-849-243-15
Query Match 3.8%; Score 44.8; DB 9; Length 3263;
Best Local Similarity 81.2%; Pred. No. 0.0087;
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 801 GCTGCTGCTGCTGCCGCGTGTGTTGCGCTGCAACTGCTGCGCCGCTGTTGTTCT 860
|||||
DB 1430 GCTGCTGCTGCTGTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371
QY 861 GCTG 864
|||||
DB 1370 GCTG 1367
RESULT 8
US-09-849-243-14/C
Sequence 14, Application US/09849243
Patent No. US20020157127A1
GENERAL INFORMATION:
APPLICANT: Kirschbaum, Bernd
Berglund, Erick
Meisterernst, Michael
Polites, Greg
TITLE OF INVENTION: PURIFICATION OF HIGHER ORDER TRANSCRIPTION
COMPLEXES FROM TRANSGENIC
NON-HUMAN ANIMALS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: HELLER, EHRMAN, WHITE & MCAULIFFE
STREET: 1666 K Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/849,243
FILING DATE: 07-May-2001
ATTORNEY/AGENT INFORMATION:
NAME: Granados, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 38005-0148
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)912-2000
TELEFAX: (202)912-2020
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4286 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: exon
LOCATION: 1..4286
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-849-243-14
Query Match 3.8%; Score 44.8; DB 9; Length 4286;
Best Local Similarity 81.2%; Pred. No. 0.01;
Matches 52; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 801 GCTGCTGCTGCTGCCGCGTGTGTTGCGCTGCAACTGCTGCGCCGCTGTTGTTCT 860
|||||
DB 2767 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708
QY 861 GCTG 864
|||||
DB 2707 GCTG 2704
RESULT 9
US-09-880-107-1748/C
Sequence 1748, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
Vockley, Joseph G.
APPLICANT: Scherf, Uwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1748
LENGTH: 6604
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 DB3783
US-09-880-107-1748
Query Match 3.8%; Score 44.8; DB 10; Length 6604;
Best Local Similarity 76.4%; Pred. No. 0.014;
Matches 55; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Downloaded from <http://ajphaphapublications.sagepub.com/> at National Archive Publishing Co on May 12, 2015

```

RESULT 1
US-09-918-995-3342
: Sequence 3342, Application US/09918995
: Publication No. US20030073623A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918,995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235,076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ. ID NOS: 38054
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3342
: LENGTH: 474
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(474)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-3342

Query Match      23.0%      Score 269; DB 9; Length 474;
Best Local Similarity 98.2%; Pred No. 5,8e-72;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0

```

Db 295 TCACCTATGCGACATTACACAGCAGCTTCAATCTCGAGTTGATCATCATGATG 354
 QY 282 TGGAGCCAGGATGATGCGGAGACATGATGAGCCTCCAGACAGTCCGCTGCATGAT 341
 Db 355 TGCAGCCAGGATGATGCGGAGATCCGTAATGACAGCTTGCAGAGACACCATGGTTGAT 414
 QY 342 CTGCTTACCTTACCGTCCAG 362
 Db 415 CTGCTTCTCTCTCAGTGCAG 435

RESULT 15

BE376590

LOCUS BE376590 663 bp mRNA linear EST 21-JUL-2000
 DEFINITION 601226419f1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3584645 5',
 mRNA sequence.

ACCESSION BE376590

VERSION BE376590.1 GI:9321955

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8743 row: f column: 06

High quality sequence stop: 498.

Location/Qualifiers

1. 663

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3584645"

/clone_lib="NCI_CGAP_Mam1"

/tissue_type="tumor, biopsy sample"

/dev_stage="10 months, virgin"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 143 a 193 c 170 g 157 t

ORIGIN

Query Match 15.4%; Score 180.2; DB 10; Length 663;

Best Local Similarity 72.6%; Pred.No. 6.6e-32;

Matches 233; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGGCTGATGAATGATGAAGGCCCCCAATGCAAGAGCTGCTGAAG 101

Db 90 CAGCTTCGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 149

QY 102 GCTCCAGGCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 161

Db 150 ACTCAGAGGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 209

QY 162 TCAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 221

Db 210 TTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269

QY 222 TCACCTCTCAGAGTACGACAGGCGGGAACCTTCACTCGAGATGATCATGACATG 281

Db 270 TCACCTATGCGACATTACACAGCAGCTTCAATCTCGAGTTGATCATCATGATG 329
 QY 282 TGGAGCCAGGATGATGCGGAGACATGATGAGCCTCCAGACAGTCCGCTGCATGAT 341
 Db 355 TGCAGCCAGGATGATGCGGAGATCCGTAATGACAGCTTGCAGAGACACCATGGTTGAT 389
 QY 342 CTGCTTACCTTACCGTCCAG 362
 Db 390 CTGCTTCTCTCTCAGTGCAG 410

Search completed: April 29, 2003, 09:24:15
 Job time : 1902.29 secs

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1AM10260 row: n column: 06
 High quality sequence stop: 628.

FEATURES

SOURCE

Location/Qualifiers
 1. 931
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="4459973"
 /clone_id="NCLCGAP.Maml"
 /tissue_type="tumor; biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: Salt;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 233 a 277 c 242 g 179 t
 ORIGIN

Query Match 15 6%; Score 181.8; DB 12; Length 931;
 Best Local Similarity 72.9%; Pred. No. 2.9e-32;

Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

42 CCGGTCCTGGGTCTGTATGATGATAGAGGCCCCCAAAATGCAAGAGCTCTGAAG 101
 156 CAGCTTCGGATCCAGTTATCAGATCATAGAGTCTCTGCAATGTAACAGTCTTAAG 215
 102 GCTCCAGGCTCGTTAACTGACAGCTGTCACAGGGCTGGAAGCTCATGCTGGGCTC 161
 216 ACTGAGGGGCTCACTTCACTGACCTGCTGACGCTGGAAGCTTCATGATGAGCTC 275
 162 TCAGTACATGTTGGTCTGTAAGCTGAGGCGCCATGAGCCATCACCAGACCGCT 221
 276 TTACCAAAATGGTGTCTGATGCTCAACCAAGAGCCATCATCACCAGACCGCT 335
 222 TCACCTCTCAGAGGTACGACGAGGCGGGAAGTCACTCACTCGAGATGATCATCAG 281
 336 TCACCTATGCGCATTCACAGCAGCTGATCATCTCGAGTGGATGATCATCAGATG 395
 282 TGAGGCCAGTGTGGGGAATCATGATGAGCTCCAGACAGTCCGCTTCATGAT 341
 396 TGAGGCCAGTGTGGGGAATCATGATGAGCTCCAGACAGTCCGCTTCATGAT 455
 342 CTGCTTACTTACCGTTCAG 362
 456 CTGCTTCTCTCTGATGATG 476

RESULT 13
 AK008060 1552 bp mRNA linear HTC 19-JAN-2002
 LOCUS
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
 enriched library, clone:2010003D20:homolog to IGSP5 PROTEIN
 (FRAME+), full insert sequence.

ACCESSION
 VERSION AK008060
 KEYWORDS AK008060.1 GI:12842009
 SOURCE HTC; CAP trapper.
 Mus musculus (Strain: C57BL/6J) adult male small intestine cDNA to
 mRNA, clone: lib: RIKEN full-length enriched mouse cDNA library
 clone: 2010003D20.

ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253

PUBMED
 REFERENCE
 AUTHORS

10349636
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

20499374
 11042159

TITLE

JOURNAL
 MEDLINE
 PUBMED

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE
 AUTHORS

4
 11076861

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
 Salto, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Salto, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schmitt, L.M., Staib, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamita, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, D., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.
 and Hayashizaki, Y.

TITLE
 JOURNAL
 MEDLINE
 PUBMED

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

REFERENCE
 AUTHORS

21085660
 11217851

5 (bases 1 to 1552)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Oho, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Salto, H., Salto, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,
 Schmitt, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.

TITLE
 JOURNAL

Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken

MGI:427273

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 454.

Location/Qualifiers

1. 471

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_1lb="Soares mouse NML"

/tissue_type="Liver"

/lab_host="DH10B"

/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTCAAGTGGAGGCGGCGGAACTTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 118 a 126 c 122 g 105 t
ORIGIN

Query Match 15.6%; Score 181.8; DB 9; Length 471;
Best Local Similarity 72.9%; Pred. No. 2.7e-32;
Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGTCTGTGTAATGATGATGAGAGGCCCCCAATGCAAGAGCTCTGAAG 101
D 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
D 115 CAGCTTCGGATCCAGTTATCATGATCATGAAAGCTCCTCAGAAATGCAAGCTTAAG 174
QY 102 GCTCCAGGCTGCTGCTCACTGACGCTCCAGGCGTGGAGCTCATGATGAGGCTC 161
D 175 ACTGAGAGGCTCACTTCACTGACGCTGACTGAGCTGGAAGCTTCTCATGTGGACTC 234
QY 162 TCAGTGACATGTTGTGCTAAGCTTCAAGGCCCATGAGCCCATCATGACCAATGACCGCT 221
D 235 TTAACCAAAATGTTGTGTGATGCTCAACCAAGGCCCATCATGACCAACACCGCT 294
QY 222 TCACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTGAGAGATGATCATCCACATG 281
D 295 TCACCTATGACGATTAACAGCACTGACATCTTCACTGAGATGATCATCCATGATG 354
QY 282 TGGAGCCAGTGTATGCGGGAACATGACATGACGCTCCAGACAGTCCGCTGCATGAT 341
D 355 TGCAGCCAGTGTACTCGGAGATCCCTGCAATGACAGCTGCAAGACGACATGGGTTGGAT 414
QY 342 CTGCTTACCTTACCGTCCAG 362
D 415 CTGCTTCTCTCATGAGTCAAG 435

RESULT 11
AM990468 611 bp mRNA linear EST 02-JUN-2000
LOCUS uc37405.y1 Soares mammary gland NMLNG Mus musculus cDNA clone

DEFINITION IMAGE:1513545 5' similar to TR:060962 060962 L549.1.; mRNA
sequence.

ACCESSION AM990468
VERSION AM990468.1 GI:8186027
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 611)

AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:940397

Seq primer: -40RP from Gibco
High quality sequence stop: 474.

Location/Qualifiers

1. 611

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_1lb="Soares mammary gland NMLNG"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 152 a 155 c 156 g 148 t
ORIGIN

Query Match 15.6%; Score 181.8; DB 10; Length 611;
Best Local Similarity 72.9%; Pred. No. 2.8e-32;
Matches 234; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 42 CCGGTTCTGGTCTGTGTAATGATGATGAGAGGCCCCCAATGCAAGAGCTCTGAAG 101
D 124 CAGCTTCGGATCCAGTTATCATGATCATGAAAGCTCCTCAGAAATGCAAGCTTAAG 183
QY 102 GCTCCAGGCTGCTGCTCACTGACGCTCCAGGCGTGGAGCTCATGATGAGGCTC 161
D 184 ACTGAGAGGCTCACTTCACTGACGCTGACTGAGCTGGAAGCTTCTCATGTGGACTC 243
QY 162 TCAGTGACATGTTGTGCTAAGCTTCAAGGCCCATGAGCCCATCATGACCAATGACCGCT 221
D 244 TTAACCAAAATGTTGTGTGATGCTCAACCAAGGCCCATCATGACCAACACCGCT 303
QY 222 TCACCTCTCAGAGTACGACGAGGCGGGAATCTTCACTGAGAGATGATCATCCACATG 281
D 304 TCACCTATGACGATTAACAGCACTGACATCTTCACTGAGATGATCATCCATGATG 363
QY 282 TGGAGCCAGTGTATGCGGGAACATGACATGACGCTCCAGACAGTCCGCTGCATGAT 341
D 364 TGCAGCCAGTGTACTCGGAGATCCCTGCAATGACAGCTGCAAGACGACATGGGTTGGAT 423
QY 342 CTGCTTACCTTACCGTCCAG 362
D 424 CTGCTTCTCTCATGAGTCAAG 444

RESULT 12

BG173684 931 bp mRNA linear EST 06-FEB-2001
LOCUS 602336823F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:445973 5',
DEFINITION mRNA sequence.

ACCESSION BG173684
VERSION BG173684.1 GI:12680387
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 931)

AUTHORS NIH-MGC http://mhc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseste@atson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

DOENY, SANDY CO INDC NATURAL RESOURCES CENTER
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smlt@email.marc.usda.gov

AUTHORS Somselgardy, J.S., Capuco, A.V., Van Vasselt, C.F., Ashmole, M.S. and Wells, K.D.

TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library

JOURNAL Unpublished (2000)

TITLE	Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library
JOURNAL	unpublished (2000)
COMMENT	Contact: Sonstegard TS

Db	628	C- -GGGCACTTATGAGGGGTCTGAAGTGTGCTGGAAGTCA-CACACTTGGGCAACGGGAGCT	684
Oy	562	TGACTTTCGCTGGGCTACCTGGAAAGAGCCTTGAAAGCCCGCAAGTCT	625
Db	685	TGACTTGTGTGGCAAGAGCTGAAAGACTTGTGCAGGCGCACAGACTCT	728
RESULT 3			
Bg740428			
LOCUS	Bg740428	725 bp	mrna
DEFINITION	602638181F1 NCI_GCAP_skn3 Homo sapiens cDNA clone IMAGE:4778789 5',	linear	EST 15-MAY-2001
VERSION	BG740428		
KEYWORDS	BG740428.1	GI:14051081	
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 725)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MSC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LRAM10635 row: b column: 06
High quality sequence stop: 725.
Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4778789"
/clone_1lb="NCI CGAP_skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."

```

Query Match	22.9%	Score 267.4	DB 12	Length 725
Best Local Similarity	99.6%	Pred. NO. 2e-52		
Matches 268	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 899	AAGAAATCTGAAAAAGACAGACAAACAAAGAACTGAGACGAAAGGGAATGAAAC			958
Db 377	AGGAATCTGAAAAAGAAAGACAAACAAAGAACTGAGACGAAAGTGAATGAAAC			436
QY 959	TCCGGCTCAATTCAGATGAACAAAAGACCAAGACACCGCTTCCTCCTCCCAATCC			1018
Db 437	TCCGGCTCAATTCAGATGAACAAAAGACCAAGACACCGCTTCCTCCTCCCAATCC			496
QY 1019	TGTGATTCACGTGATCTCGAACAAAAGAAACAGTAGTCTTGCGCCTCTCCACACGGGGCT			1078
Db 497	TGTGATTCACGTGATCTCGAACAAAAGAAACAGTAGTCTTGCGCCTCTCCACACGGGGCT			556
QY 1079	GATCAAGTCCACCCAGGCCACGACGATCATCCACAGGCTTCCTTTAACTGGCCAGTCTT			1138
Db 557	GATCAAGTCCACCCAGGCCACGACGATCATCCACAGGCTTCCTTTAACTGGCCAGTCTT			616
QY 1139	GAGAAGTCAGTAATATACAACGTGTAGTATA			1167
Db 617	GAGAAGTCAGTAATATACAACGTGTAGTATA			645

RESULT 4	
LOCUS	BB663870
DEFINITION	BB663870 RIKEN full-length enriched, 0 day neonate lung Mus
ACCESSION	BB663870
VERSION	BB663870.1 GI:16497624
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 631)
AUTHORS	Arakawa,T., Carandini,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,T., Itoh,M., Kawai,T., Kono,T., Kurauchi,T.,

CONTACT: Yoshinobu Hayashizaki
Contact: yoshinobu.hayashizaki@riken.jp
RIKEN MOUSE ESTS (Arakawa, T., et al. 2001)
TITLE *Genetic analysis of the mouse genome*
JOURNAL *Genetics*
COMMENT

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagl, K., Fujiwake, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Okawa, K., Tanaka, T., Matsunura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamane, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

```

1. 631
SOURCE
1. 631
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E030016M12"
/clone_lib="RIKEN full-length enriched, 0 day neonate
lung"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/notes="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was

```


source 1. 784
 /organism="Mus musculus"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:4917803"
 /clone_id="NCI_CGAP_Mam4"
 /tissue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 188 a 202 c 208 g 186 t

ORIGIN

Query Match 26.2%; Score 306; DB 12; Length 784;
 Best Local Similarity 70.6%; Pred. No. 1.7e-61;
 Matches 436; Conservative 0; Mismatches 180; Indels 2; Gaps 2;

42 CCGGTTCTGGGTCGTGAATGATAGAGGCCCCCAATGCAAGTCTGAAG 101
 148 CAGCTTCGGATCGATTACATGATAGAGGCTCTCAGATGTAACGCTTAAG 207
 102 GCTCCAGGCTCGCTTCACCTGCAAGGCTCCAGGCTGGAAGCTCATCTATG 161
 208 ACTCAGAGGCTCACTTCACTGCAAGGCTGCAAGGCTGGAAGCTTCTATG 267
 162 TCAGTGCATGCTGCTCTCAAGGCTCAGGCTGCAAGGCTGCAAGGCTGCA 221
 268 TTAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
 222 TCACCTCTCAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 281
 328 TCACCTCTCAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387
 282 TGAGAGCCAGTATGTCGGGAGATCAGATGAGCTCCAGAACAGTCCGCTG 341
 388 TGAGAGCCAGTATGTCGGGAGATCAGATGAGCTCCAGAACAGTCCGCTG 447
 342 CTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 401
 448 CTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 507
 402 TCGGTGAGTAACTTGAAGTATGATGATGATGATGATGATGATGATGAT 461
 508 TCACCTGAGGAGTAACTTGAAGTATGATGATGATGATGATGATGATGAT 567
 462 ATATTTCCTGGAGCTCGCTCTCTGTCAGCCTTCAAGCTATTTATTTG 521
 568 ATATTTCCTGGAGCTCGCTCTCTGTCAGCCTTCAAGCTATTTATTTG 627
 522 CCAGGACCTTCAAGTGCAGTGCATCTGCTGCTGCTGCTGCTGCTGCTG 581
 628 CGGACCACTTATGAGGCTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTG 686
 582 TGACTTCGTCGTCAGTGCAGTGCATCTGCTGCTGCTGCTGCTGCTGCTG 641
 687 TGACTTCGTCGTCAGTGCAGTGCATCTGCTGCTGCTGCTGCTGCTGCTG 745
 642 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 746 CTGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763

RESULT 2
 BF784177 794 bp mRNA linear EST 12-JAN-2001
 LOCUS BF784177
 DEFINITION 602108039F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4236489
 5', mRNA sequence.
 ACCESSION BF784177

VERSION BF784177.1 GI:12089213
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 794)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE NIH-MGC
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN, at: <http://image.llnl.gov>
 Plate: L1AM9844 row: n column: 10
 High quality sequence stop: 694.
 Location/Qualifiers
 1. 794
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:4236489"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-Sport6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library." 1

BASE COUNT 186 a 214 c 215 g 179 t

ORIGIN

Query Match 23.7%; Score 276.4; DB 12; Length 794;
 Best Local Similarity 69.3%; Pred. No. 1.6e-54;
 Matches 405; Conservative 0; Mismatches 176; Indels 3; Gaps 2;

42 CCGGTTCTGGGTCGTGAATGATAGAGGCCCCCAATGCAAGTCTGAAG 101
 148 CAGCTTCGGATCGATTACATGATAGAGGCTCTCAGATGTAACGCTTAAG 207
 102 GCTCCAGGCTCGCTTCACCTGCAAGGCTCCAGGCTGGAAGCTCATCTATG 161
 208 ACTCAGAGGCTCACTTCACTGCAAGGCTGCAAGGCTGGAAGCTTCTATG 267
 162 TCAGTGCATGCTGCTCTCAAGGCTCAGGCTGCAAGGCTGCAAGGCTGCA 221
 268 TTAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
 222 TCACCTCTCAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 281
 328 TCACCTCTCAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 387
 282 TGAGAGCCAGTATGTCGGGAGATCAGATGAGCTCCAGAACAGTCCGCTG 341
 388 TGAGAGCCAGTATGTCGGGAGATCAGATGAGCTCCAGAACAGTCCGCTG 447
 342 CTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 401
 448 CTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 507
 402 TCGGTGAGTAACTTGAAGTATGATGATGATGATGATGATGATGATGAT 461
 508 TCACCTGAGGAGTAACTTGAAGTATGATGATGATGATGATGATGATGAT 567
 462 ATATTTCCTGGAGCTCGCTCTCTGTCAGCCTTCAAGCTATTTATTTG 521
 568 ATATTTCCTGGAGCTCGCTCTCTGTCAGCCTTCAAGCTATTTATTTG 627
 522 CCAGGACCTTCAAGTGCAGTGCATCTGCTGCTGCTGCTGCTGCTGCTG 581

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 23:25:02 ; Search time 1894.29 Seconds
(without alignments)
9985.963 Million cell updates/sec

Title: US-09-729-264-3

Perfect score: 1168
Sequence: 1 agtgcacatgcgtgcgcagag.....gtatcacacatgtagatag 1168

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues.

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	306	26.2	784	12	BC683468 602796941
2	276.4	23.7	794	12	BF784177 602108039
3	267.4	22.9	725	12	BF740428 602633817
4	256.8	22.0	631	10	BB63870 BB63870
5	242	20.7	527	10	BE032610 132035 MA
6	193.2	16.5	517	10	BE476432 159670 BA

7	190.8	16.3	622	17	AG069679	AG069679 Pan trogl
8	190.8	15.3	577	17	AG107877	AG107877 Pan trogl
9	185.8	15.9	524	17	AV601192	AV601192 AV601192
10	181.8	15.6	471	9	AA265274	AA265274 mx91d09.r
11	181.8	15.6	611	10	AM990468	AM990468 uc37d05.y
12	181.8	15.6	931	12	BG173684	BG173684 602336823
13	181.8	15.5	1552	11	AK008060	AK008060 Mus muscu
14	180.8	15.6	486	9	AI425263	AI425263 mx51d09.y
15	180.2	15.4	663	10	BE376590	BE376590 601226419
16	178.6	15.3	650	10	BE625217	BE625217 BB625217
17	165	14.1	600	17	AE379623	AE379623 IM0134622
18	153.2	13.1	216	12	BG206666	BG206666 RST6117
19	149	12.8	556	16	AI788300	AI788300 uk56d01.y
20	147.4	12.6	530	9	AI790785	AI790785 uk28a12.y
21	145.4	12.4	291	17	AZ411779	AZ411779 IM0185B04
22	141.6	12.1	423	10	BB846577	BB846577 BB846577
23	128.8	11.0	471	12	BF040046	BF040046 BP250004B
24	124	10.6	685	17	AG142221	AG142221 Pan trogl
25	123.2	10.5	394	10	BB846133	BB846133 BB846133
26	110	9.4	785	13	BI452873	BI452873 603170211
27	102.8	8.8	889	13	BI454276	BI454276 603171509
28	100	8.6	185	14	N47851	N47851 yw95h05.r1
29	100	8.6	234	14	N93995	N93995 za65f09.r1
30	98.4	8.4	525	17	AQ403719	AQ403719 HS-5049_A
31	78.6	6.7	360	9	AI647601	AI647601 uk41f12.x
32	78.6	6.7	374	9	AI648720	AI648720 uk28a12.x
33	78.6	6.7	406	9	AI788973	AI788973 uk56d01.x
34	77	6.6	644	10	BB531388	BB531388 BB531388
35	67.8	5.8	397	10	AM112084	AM112084 KC7851 mo
36	62.6	5.4	282	10	BB564363	BB564363 BB564363
37	56	4.8	818	12	BG445021	BG445021 GA_Ea002
38	52.6	4.5	422	9	AI756062	AI756062 ETEESTa09
39	50.8	4.3	458	9	AL514929	AL514929 AL514929
40	50	4.3	447	10	AV431233	AV431233 AV431233
41	49.8	4.3	428	9	AA695064	AA695064 GM02059.5
42	49.6	4.2	925	17	CNS0381E	AL232043 Tetradon
43	49.6	4.2	925	17	CNS0381E	AL065607 Drosophil
44	49.4	4.2	481	10	BE017134	BE017134 FK76f06.y
45	48.6	4.2	567	14	BQ598989	BQ598989 MT-P-E6-4

ALIGNMENTS

RESULT 1
LOCUS BC683468 784 bp mRNA linear EST 29-MAY-2001
DEFINITION 602796941F1 NCI-CGAP_Mam4 Mus musculus cDNA clone IMAGE:4917803 5',
mRNA sequence.
ACCESSION BC683468
VERSION BC683468.1 GI:14214006
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 784)
AUTHORS NIH-MGC
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL Unpublished (1999)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM10830 row: b column: 12
High quality sequence stop: 768.
Location/Qualifiers

FEATURES

Wed Apr 30 10:00:03 2003

us-09-729-264-1.rge

Page 17

||||| | ||||| |||
Db 121168 AGAGGCTCCTCAACTGCACTGTGACTCATGGCCTGA 121205

Search completed: April 29, 2003, 14:10:30
Job time : 4555.27 secs


```

repeat_region      33468..33678
                    /evidence=not_experimental
                    /rpt_family="THelB"
repeat_region      34563..34620
                    /note="homology = 93.10%, score = 21, counts = 2"
                    /evidence=not_experimental
                    /rpt_type=tandem
                    /rpt_unit=tttcagaaagatacgtacttaggca
repeat_region      complement(35785..35824)
                    /evidence=not_experimental
                    /rpt_family="MLT2CB"
repeat_region      complement(35924..36345)
                    /evidence=not_experimental
                    /rpt_family="MLT2CB"
repeat_region      complement(36403..36475)
                    /evidence=not_experimental
                    /rpt_family="MLT2CB"
repeat_region      complement(36717..36828)
                    /evidence=not_experimental
                    /rpt_family="MER67C"
repeat_region      complement(37064..37364)
                    /evidence=not_experimental
                    /rpt_family="AluY"
repeat_region      37440..37501
                    /note="homology = 98.40%, score = 29, counts = 2"
                    /evidence=not_experimental
                    /rpt_type=tandem
                    /rpt_unit=aaaaccttcactgtcttcaaacctgctc
repeat_region      37518..38158
                    /evidence=not_experimental
                    /rpt_family="MLT1D"
repeat_region      39884..40122
                    /evidence=not_experimental
                    /rpt_family="MLT1D"
repeat_region      complement(40243..40542)
                    /evidence=not_experimental
                    /rpt_family="AluSx"
repeat_region      40543..41180
                    /evidence=not_experimental
                    /rpt_family="L2"
repeat_region      complement(42066..42114)
                    /evidence=not_experimental
                    /rpt_family="L2"
repeat_region      complement(42317..42317)
                    /evidence=not_experimental
                    /rpt_family="MIR"
repeat_region      complement(42332..42710)
                    /evidence=not_experimental
                    /rpt_family="L1B1"
repeat_region      complement(42758..46751)
                    /evidence=not_experimental
                    /rpt_family="L1B1"
repeat_region      complement(46792..47084)
                    /evidence=not_experimental
                    /rpt_family="L1MA"
repeat_region      47189..47992
                    /evidence=not_experimental
                    /rpt_family="L2"
repeat_region      complement(48002..48132)
                    /evidence=not_experimental
                    /rpt_family="MIR"
repeat_region      48905..49310

```

Query Match 8.5%; Score 100; DB 9; Length 199665;
 Best Local Similarity 100.0%; Pred. No. 3.3e-13;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1076 CCAGCGGGGATCAGTCGCAACCCAGGCGAGAGTCAATCCAGAGGCTTTTAACT 1135
      |||||||
Db 17555 CCAGCGGGGCTGATCAGACGTCACCCAGGCGAGCAAGTCAATCCAGAGGCTTTTAACT 17614
      |||||||
QY 1136 GGCGACTCCTGAGAAAGTCAGTAATCAACTGATAG 1175
      |||||||
Db 17615 GGCGACTCCTGAGAAAGTCAGTAATCAACTGATAG 17654
      |||||||

RESULT 15
AC120346
LOCUS
DEFINITION
AC120346 175861 bp DNA linear HTG 09-MAY-2002
Mus musculus clone RP23-147E11, WORKING DRAFT SEQUENCE, 22 ordered
pieces.
AC120346
AC120346.3 GI:20514894
VERSION
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
MUS MUSCULUS.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 175861)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-147E11
Unpublished
2 (bases 1 to 175861)

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
JOURNAL
AUTHORS
REFERENCE
AUTHORS
2 (bases 1 to 175861)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, L.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicot, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Stuhms, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 175861)

```

```

TITLE
JOURNAL
JOURNAL
AUTHORS
REFERENCE
AUTHORS
3 (bases 1 to 175861)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, L.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicot, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Stuhms, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 175861)

```


REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
1	Welcher, A.A., Sarmiento, U.M., Schultze, H.J. and Chute, H.T.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Amgen, Inc. (US)	Location/Qualifiers
source				1..895
CDS				/organism="Mus musculus" /db_xref="taxon:10090" 53..865 /note="unnamed protein product" /codon_start=1 /protein_id="CAD28221.1" /db_xref="GI:19575337" /translation="MEGSRDYLAVLYLAOLTASGSSVOIEEPONTVYLKDSFAH NCTYTHGRKLMTLNLQKRVYLSLTQGGITLNNRRTYASTVSTFSELITHDVP DSGSVQCSLONSHGSAFSLVDSIGEGPALPMALILAVFSLLLIYLII PCCCAASRREESTYQMEIRKRSANMRKMDPEFKLSGKENYGVSSDEAKMAOTS LPPLSAEVLSPLEKSSSLPYELKMHQGPATHPRVSFDIASPKVRNVTLV"
BASE COUNT	250 a	223 c	212 g	210 t
ORIGIN				
Query Match	15.8%	Score 186;	DB 6;	Length 895;
Best Local Similarity	71.1%;	Pred. No. 4.4e-34;		
Matches 246;	Conservative 0;	Mismatches 100;	Indels 0;	Gaps 0;
QY	24	GAGATGGGGCTTGATTTCTCCACGGTTTGCGTGTGTAATGAAGTCAAGGC	83	
DB	83	GTGCTGTCATCTTGCTGGCTACAGTGCAGCTTCGGATCAGTTACAGTAAAGGT	142	
QY	84	CCCCGAATGCAACAGTCTCGAAGGGCTCCAGGCTCGCTTCACTGACCGCTCCAG	143	
DB	143	CCTCAGATGTAAACAGTCTTAAAGAGCTAAGAGGCTCACTTCACTGACCGTACTAC	202	
QY	144	GGCTGGAAGTCATCATGTGGGCTCTCAGTACATGTGTGTCTAAGGCTCAGGCCCATG	203	
DB	203	GCGTGAAGCTTCTCATGTGACCTTTAACCAAAAGTGGTGGCTGACGTCCACCA	262	
QY	204	GAGCCATCATCACCAATGACCGCTTCACTCTCAGAGAGTACGACGAGCGGGAACTTC	263	
DB	263	GGACCCATCATCACCAACACCGCTTCACTTACGCAATTCACACGACACTGACGCTTC	322	
QY	264	ACCTCGGATGATCATCTACCAATGTGGAGCCCACTGATTCGGGGACATCATGATGACG	323	
DB	323	ATCTCGGATGATCATCTCATGATGTGACGCCCACTGATGCTCGGATTCGTCATGACG	382	
QY	324	CTCCAGACAGTGGCTTCATGATGATCTTACCTTACCGTCCAG	369	
DB	383	CTGCAGACAGCCATGGGTTTGATGCTGCTTCTCTCAGTGCAG	428	
RESULT 12				
LOCUS	BC004806	1556 bp	mRNA	linear
DEFINITION	Mus musculus, RIKEN cDNA 2010003020 gene, clone MGC:7960			
ACCESSION	BC004806			
VERSION	BC004806.1			
KEYWORDS	MG.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 (Phases 1 to 1556)			
TITLE	Strausberg, R.			
JOURNAL	Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			

[illegible]

```

repeat_region /rpt_family="MER34"
                complement(17633..17969)
                /evidence=not_experimental
                /rpt_family="AluSg1"
                complement(18004..18334)
                /evidence=not_experimental
                /rpt_family="L1R17"
                complement(18335..18632)
                /evidence=not_experimental
                /rpt_family="HERV17"
                complement(18352..18420)
                /note="GRAIL, score = 51.000%, comment = good shadow"
                /evidence=not_experimental
exon            complement(18397..18584)
                /note="GRAIL, score = 66.000%, comment = good"
                /evidence=not_experimental
                complement(18643..24647)
                /evidence=not_experimental
repeat_region   /evidence=not_experimental
                /rpt_family="HERV17"
                complement(19228..19361)
                /note="MZF, score = 92.8%"
                /evidence=not_experimental
                20900..21135
                /note="MZF, score = 51.4%"
                /evidence=not_experimental
                21150..21275
                /note="Xpound exon prediction, score = 74% (0%)"
                /evidence=not_experimental
exon
exon

```

```

Query Match      27.3%; Score 320.2; DB 9; Length 142742;
Best Local Similarity 97.6%; Pred. No. 2,8e-66;
Matches 325; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 42 TTCCTCAGCGTCTGTGGTCTGTATGAAGTATGAAGGCCCCCGAGATGCAACAGTC 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44265 TACCTTCAGAGTCTGTGGTCTGTATGAAGTATGAAGGCCCCCGAGATGCAACAGTC 44324

QY 102 CTGAAGGGCTCCAGGCTGCTTCAACGACCGTCTCCAGGGCTGAAGCTCATCATG 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44325 CTGAAGGGCTCCAGGCTGCTTCAACGACCGTCTCCAGGGCTGAAGCTCATCATG 44384

QY 162 TGGGCTCTCGTACATGGTGGTGTCTAAGGCTGAGGCCCATGGAGCCCATTCATCCATC 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44385 TGGGCTCTCGTACATGGTGGTGTCTAAGGCTGAGGCCCATGGAGCCCATTCATCCATC 44444

QY 222 GACCGCTTCACTCTCTAGAGTACGACCAAGGGGGAAGTCACTCTCGAGATGATCATC 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44445 GACCGCTTCACTCTCTAGAGTACGACCAAGGGGGAAGTCACTCTCGAGATGATCATC 44504

QY 282 CACAATGTGAGCCCAAGTATTCGGGGAACATCATGACCTCCAGAACATGCGCTG 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44505 CACAATGTGAGCCCAAGTATTCGGGGAACATCATGACCTCCAGAACATGCGCTG 44564

QY 342 CATGATCTGCTTACCTACCTACCGTCAAGTTATG 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44565 CATGATCTGCTTACCTACCTACCGTCAAGTTATG 44597

```

```

RESULT 8
LOCUS      HS21C080          340000 bp      DNA      linear      PRI 24-MAY-2000
DEFINITION Homo sapiens chromosome 21 segment HS21C080.
ACCESSION  AL163280 AP001735 BA000005
VERSION     AL163280.2 GI:7717369
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 340000)
AUTHORS     Hattori,M., Fujiyama,A., Taylor,T.D., Matsumae,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Sasaki,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,

```

TITLE JOURNAL COMMENT

Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Kanser,J., Beck,A., Klages,S.,
Hennig,S., Resselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.

Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e-mail: sakaki@sc.riken.go.jp/
* URL: http://hgp-gsc.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e-mail: gscf-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8562, Japan,
* e-mail: shimizue@mb-med.keio.ac.jp/
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:
info.genome@bf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
* Inestrasse 73, D-14195 Berlin, Germany,
* e-mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.

FEATURES

Location/Qualifiers
1..340000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1..125946
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P160L9, 5' partial"
/clone_id="RPC11,3-5 PAC library"
/note="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/map="21q22.3"
/clone="P70124"
/clone_id="RPC11,3-5 PAC library"
/note="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_id="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000

SOURCE

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 340000)
AUTHORS Hattori,M., Fujiyama,A., Taylor,T.D., Matsumae,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Sasaki,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,

```

repeat_region complement(140. .415)
/evidence-not_experimental
/rpt_family="MLT2F"
repeat_region complement(668. .881)
/evidence-not_experimental
/rpt_family="MER58A"
complement(1078. .1162)
/note="GRAIL, score = 86.000%, comment = excellent shadow"
exon /evidence-not_experimental
1114. .1403
repeat_region /evidence-not_experimental
/rpt_family="MLT1I"
1481. .1603
repeat_region /evidence-not_experimental
/rpt_family="L2"
1576. .1710
exon /note="MZE, score = 66.1%"
/evidence-not_experimental
1576. .1710
exon /note="GRAIL, score = 99.000%, comment = excellent"
/evidence-not_experimental
complement(1693. .1741)
repeat_region /evidence-not_experimental
/rpt_family="MIR"
complement(1780. .2326)
repeat_region /evidence-not_experimental
/rpt_family="MLT1F"
complement(1834. .1955)
exon /note="GRAIL, score = 50.000%, comment = good shadow"
/evidence-not_experimental
2629. .2831
repeat_region /evidence-not_experimental
/rpt_family="AluSx"
3012. .3122
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"
3123. .3478
repeat_region /evidence-not_experimental
/rpt_family="THER1B"
3479. .3618
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"
complement(3502. .3540)
exon /note="GRAIL, score = 65.000%, comment = good shadow"
/evidence-not_experimental
4674. .4781
exon /note="GRAIL, score = 42.000%, comment = marginal"
/evidence-not_experimental
4712. .4749
exon /note="Xpound exon prediction, score = 70% (0%)"
/evidence-not_experimental
5385. .6104
repeat_region /evidence-not_experimental
/rpt_family="HERVL"
6682. .6751
exon /note="MZE, score = 95.1%"
/evidence-not_experimental
6746. .6751
exon /note="Genscan, score = 1.83%, comment = Initial_exon 6 bp
frame: 1 phase: 0"
/evidence-not_experimental
6901. .7339
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
7347. .7457
repeat_region /evidence-not_experimental
/rpt_family="LIR16C"
complement(7561. .7696)
repeat_region /evidence-not_experimental
/rpt_family="MIR"
complement(7819. .8220)
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"

exon 8097. .8119
/note="GRAIL, score = 45.000%, comment = marginal"
/evidence-not_experimental
complement(8738. .9232)
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
complement(9273. .9420)
exon /note="GRAIL, score = 71.000%, comment = good"
/evidence-not_experimental
9290. .9371
exon /note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence-not_experimental
9966. .10041
repeat_region /note="homology = 100.00%, score = 38, counts = 2"
/evidence-not_experimental
/rpt_type=" tandem
/rpt_unit="ctcttgcacattctctgttgagagtcgttatcgc
10244. .10664
repeat_region /evidence-not_experimental
/rpt_family="MLT1E"
10665. .11028
repeat_region /evidence-not_experimental
/rpt_family="MLT1F"
11318. .11417
repeat_region /evidence-not_experimental
/rpt_family="L1MC/D"
11503. .11818
repeat_region /evidence-not_experimental
/rpt_family="AluSg"
complement(11525. .11777)
exon /note="MZE, score = 79.4%"
/evidence-not_experimental
complement(12346. .12827)
repeat_region /evidence-not_experimental
/rpt_family="MLT1C"
12902. .13180
repeat_region /evidence-not_experimental
/rpt_family="AluSg"
complement(13057. .13161)
exon /note="MZE, score = 76.1%"
/evidence-not_experimental
complement(13226. .13364)
repeat_region /evidence-not_experimental
/rpt_family="L1MC2"
complement(13613. .13908)
repeat_region /evidence-not_experimental
/rpt_family="AluSx"
complement(14317. .14614)
exon /evidence-not_experimental
/rpt_family="LIR37A"
complement(14962. .15036)
exon /note="MZE, score = 93%"
/evidence-not_experimental
15255. .15299
repeat_region /evidence-not_experimental
/rpt_family="L1P6"
complement(15459. .15497)
exon /note="Xpound exon prediction, score = 62% (0%)"
/evidence-not_experimental
16045. .16768
repeat_region /evidence-not_experimental
/rpt_family="MER21A"
complement(16555. .16655)
exon /note="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
16823. .16850
repeat_region /evidence-not_experimental
/rpt_family="MER34"
complement(16851. .17208)
repeat_region /evidence-not_experimental
/rpt_family="THER1A"
17209. .17632
repeat_region /evidence-not_experimental

```

```

LOCUS       AX380404                1195 bp    DNA             linear      PAT 18-MAR-2002
DEFINITION   Sequence 9 from Patent WO0200710.
ACCESSION    AX380404
VERSION      AX380404.1  GI:19575334
KEYWORDS
SOURCE
  ORGANISM    Mus musculus.
               house mouse.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
  1. Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
  B7-like molecules and uses thereof
  Patent: WO 0200710-A 9 03-JAN-2002;
  Angen, Inc. (US)

FEATURES
  source
    1..1195
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    53..1165
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD28220.1"
    /db_xref="GI:19575335"
    /translation="MEGSMRDYLAIVLIIAQLTASGSYQIIEGPNVTVLIDSEAHF
    NCTVHGMKILMTLNMVYLSITGPIITNNRPITYASYNSTDFISLIIHDVPS
    DSGVQCSLQNSHSGESAFSLVQVMTLIPSNLIVTGEPCNVCVAVGWSLPDI
    SWELEVPVSHSNISFLEPGENRVLVDLPPLGLTGEPCNVCVAVGWSLPDI
    TVVQPPDSISGEGRPALPTWAILLAVASFLILIVLIIIFCCCSARRESESTY
    ONEIRKSAMNRITKADPETKLSKENYSSDEAKQFASLIPRSASVLSPEKRS
    SLPEYELNKHQPPATHPRVSPDIASPVQAVRVNTLV"

BASE COUNT   313 a      304 c      291 g      287 t

Query Match      32.4%; Score 381; DB 6; Length 1195;
Best Local Similarity 61.9%; Pred. No. 4.6e-81;
Matches 712; Conservative 0; Mismatches 370; Indels 69; Gaps 4;

24  GAGATGGGGCTTGATTTCTCCACAGCTTGGGCTGTGATAGTACATAGAGGC 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83  GTGCTGTCTATCTCGCTCAGCTGACAGCTTCGGATCCGATTATCATAGATAGAGGT 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84  CCCAGAAATCAACAGCTGAGAGGGCTCCAGGCTCCGCTCAACTGACCGCTCCAG 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
143  CCTCAGATTAACAGCTTAAGAGCTCAGAGGCTCATTCACTGACCGTGGACTCAC 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
144  GGCTGAGCTCATCATGAGGCTCTCAGTACATGAGTGTGCTAAGCGTACAGGCCATG 203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
203  GGCTGAGGCTCTCAGTACATGAGTGTGCTAAGCGTACAGGCCATG 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
204  GAGCCCATCATCAACAGTACCGCTCAGCTCAGAGGACAGCAGGCGGGAACCTTC 263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263  GGACCCATCATCAACAGTACCGCTCAGCTCAGAGGACAGCAGTGCACACTTC 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264  ACCTGAGATGATATCCATCCAGTGTGAGCCAGTATGCGGGAACATACATGACAG 323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
323  ATCTCGAGATGATATCCATCCAGTGTGAGCCAGTATGCGGGAACATACATGACAG 382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324  CTCGAGAACAGTGGCTGATGATGCTTACCTTACCGTCCAGTATAGGAGAGCTG 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
383  CTCGAGAACAGTGGCTGATGATGCTTACCTTACCGTCCAGTATAGGAGAGCTG 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384  TTGATTTCCAGTGTATCTTGATGCGCTGAGATGAGAACCTTGTAACTTACTGTCTA 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443  AACATTTCTTACGACAACTTATAGTACAGGAGTGAACCCCTGATATGACTTGTCTAT 502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
444  CCTCAGACTGAGCCGGGATTTCTCGGAGAGTGGTCTCTCTGCTCAGACCAT 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
503  GCGTGGGCTGAGCTCAGCTACCTCCGGATATTTCTCGGAGAGTGGTCTCTCTGCTCAGACCAT 562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
504  TCAGATATATTTTGTTCGGAGGCCAGGACCTTCAAAAGTGCAGTGCATCTGGCT 563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
563  TCAGATATATTTTGTTCGGAGGCCAGGACCTTCAAAAGTGCAGTGCATCTGGCT 622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 564 CTGACCCACAGACGATGAGCTTGAAGTCTGAGCTACCTGAGAGAGCCTGAGGCC 623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 623 CTCACACACACTGGGACAGGAGACCTTGAAGTCTGAGAGCTGAGAGCTTGCAGGCC 682
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 624 CGCAAGCTGCACTGTAAATCTCACTGTGATGCGTGTCCCAAGACAGTGGAGTGT 683
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 683 AGCAAGCTCTTCACTGTCAACCTGACTGTGTTGACGCTCCAGTCAAGATTTGGAGAG 742
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 684 ATTAATATCCAGGATATATCAAGTTTACGAGTTTATGTTTCTGCTGCTGCTG 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 743 GA-----AGGCCAGCACTGCCGACCTGG 766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 744 GCGCAAGTTGAGCTTGGACTAGACGACACCATCTTGTGACGCCGAGTACTCTTACA 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 767 GCATCATCTCTGCTGACATGCGCTTTCCTGCTGCTGATCTGATCTGATTTGAT 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 804 ATACGCTGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 827 ATTAATATCTGTTGCTGTTGT-----GCTTCAGAGAGAAAGAGAAATC 873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 864 TTCTGCTGAGAGAGAGAGATTTGATTTCAATTTCAAAAGAAATCGAAGAAAGAG 923
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 874 TACTTATCAAAATGAAATAGAGAAATCTGCAACATGAGAGCA-----AC 919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 924 AAGACAAACAAAGAACTGACAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 983
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 920 AAGACAGATCCGAGACAAAGTTAAAGTGAAGAGAAAGTGAAGTGAAGTGAAGT 979
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 984 GAACAAAGACACAGACACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1043
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 980 GAGGCAAGGCTGACAGACATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1039
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1044 GAACAAAGAAAGATGATGTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1103
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1040 GAAAGAGAGAGAGAGT-----GCTTCTCTATTCAGAGAACTCAATTAACATCAGCCGGT 1093
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1104 CCAGCAAGTATCATCAAGGCTTCTTAAATCTGCGCACTCTGAGAGGTCAGTAATACA 1163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1094 CCAGCAAGTATCATCAAGGCTTCTTAAATCTGCGCACTCTGAGAGGTCAGTAATGTG 1153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1164 ACTGATGATA 1174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1154 ACTTATGATGA 1164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
AF121782
LOCUS       AF121782                142742 bp    DNA             linear      PRI 02-FEB-1999
DEFINITION   Homo sapiens chromosome 21q22.3 PAC 206A10, complete sequence.
ACCESSION    AF121782
VERSION      AF121782.1  GI:4210991
KEYWORDS
  SOURCE
    Homo sapiens.
    Homo sapiens.
  ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  Taudien, S., Dagand, E., Hildmann, T., Nordstiek, G., Drescher, B.,
  Schattevoy, R., Weber, J., Schilling, M., Menzel, U., Vespö, M.L. and
  Rosenhal, A.

REFERENCE
  1 (bases 1 to 142742)
  Taudien, S., Dagand, E., Hildmann, T., Nordstiek, G., Drescher, B.,
  Schattevoy, R., Weber, J., Schilling, M., Menzel, U., Vespö, M.L. and
  Rosenhal, A.

AUTHORS
  Taudien, S., Dagand, E., Hildmann, T., Nordstiek, G., Drescher, B.,
  Schattevoy, R., Weber, J., Schilling, M., Menzel, U., Vespö, M.L. and
  Rosenhal, A.

TITLE
  Direct Submission
  Submitted (21-JUN-1999) Genome Analysis, Institute for Molecular
  Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

JOURNAL
  Location/Qualifiers
    1..142742
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="PAC 206A10"
    3..118
    /evidence=not experimental
    /rpt_family="MIR"

```


QY 351 GCTTACCTTACGCTCCAGTATGAGAGCTGTTCATTCCAGTGTATCTTGTAGTC 410
 |||||||
 DB 416 GCTTACCTTACGCTCCAGTATGAGAGCTGTTCATTCCAGTGTATCTTGTAGTC 475
 |||||||
 QY 411 GCTGAAATGAACCTTGTAGTACTTGTCTTACCTCTACCTGAGCCGGCTCCGGAT 470
 |||||||
 DB 476 GCTGAAATGAACCTTGTAGTACTTGTCTTACCTCTACCTGAGCCGGCTCCGGAT 535
 |||||||
 QY 471 ATTTCCTGGAGCTGGTCTCTGTCAGCCATTCAGCTATTTATTTTCTCCGAGCC 530
 |||||||
 DB 536 ATTTCCTGGAGCTGGTCTCTGTCAGCCATTCAGCTATTTATTTTCTCCGAGCC 595
 |||||||
 QY 531 AGGACCTTCAAAAGTSCAGTGAAGCATCTGGCTCTGACCCACAGACAAGTGGAGCTTTG 590
 |||||||
 DB 596 AGGACCTTCAAAAGTSCAGTGAAGCATCTGGCTCTGACCCACAGACAAGTGGAGCTTTG 655
 |||||||
 QY 591 ACTTGGGTGCTCTCTGAGAAAGCTGGAAGCCCGCAATCTGCACTTAATCTCACT 650
 |||||||
 DB 656 ACTTGGGTGCTCTCTGAGAAAGCTGGAAGCCCGCAATCTGCACTTAATCTCACT 715
 |||||||
 QY 651 GTGATTCGGTCTCCCAAGACACTGAGGCTGATTAATATTCAGCTGATTAATCACT 710
 |||||||
 DB 716 GTGATTCGGTCTCCCAAGACACTGAGGCTGATTAATATTCAGCTGATTAATCACT 775
 |||||||
 QY 711 TTACCGAGTTTAAAGTTTTCATTGCTTACTTGGGCAAAAGTTGAGCTTGGAGCTAGCAGC 770
 |||||||
 DB 776 TTACCGAGTTTAAAGTTTTCATTGCTTACTTGGGCAAAAGTTGAGCTTGGAGCTAGCAGC 835
 |||||||
 QY 771 ACCATGCTTTCAGCGCCGAGCTGCTTACATACGCTGCTGCTGCTGCTGCTGCT 830
 |||||||
 DB 836 ACCATGCTTTCAGCGCCGAGCTGCTTACATACGCTGCTGCTGCTGCTGCTGCTGCT 895
 |||||||
 QY 831 TGTGTGAGTGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
 |||||||
 DB 896 TGTGTGAGTGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 |||||||
 QY 891 CGTATTCATTTTCAAAAGAAATCTGAAAAAGAGAGACAACAAGAACTGAGACAGAA 950
 |||||||
 DB 956 CGTATTCATTTTCAAAAGAAATCTGAAAAAGAGAGACAACAAGAACTGAGACAGAA 1015
 |||||||
 QY 951 AGTGAATGAATGAATCCGGCTACATTCAGATGAACAAAGAGCCAGACACCGCTTCT 1010
 |||||||
 DB 1016 AGTGAATGAATGAATCCGGCTACATTCAGATGAACAAAGAGCCAGACACCGCTTCT 1075
 |||||||
 QY 1011 CTCCTCTCCCAATCTCTGTGATCTCACTGATCTGAGCAAAAGAAAGTGTGCTGCT 1070
 |||||||
 DB 1076 CTCCTCTCCCAATCTCTGTGATCTCACTGATCTGAGCAAAAGAAAGTGTGCTGCT 1135
 |||||||
 QY 1071 CTCCTCTCCCAATCTCTGTGATCTCACTGATCTGAGCAAAAGAAAGTGTGCTGCT 1130
 |||||||
 DB 1136 CTCCTCTCCCAATCTCTGTGATCTCACTGATCTGAGCAAAAGAAAGTGTGCTGCT 1195
 |||||||
 QY 1131 ATCTGGCCAGCTCTGAGAAAGCTGATTAATACAGCTGTGATATAG 1175
 |||||||
 DB 1196 ATCTGGCCAGCTCTGAGAAAGCTGATTAATACAGCTGTGATATAG 1240
 |||||||

RESULT 4
 AX380398 1168 bp DNA linear PAT 18-MAR-2002
 LOCUS AX380398
 DEFINITION Sequence 3 from Patent WO0200710.
 ACCESSION AX380398
 VERSION AX380398.1 GI:19575328
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
 AUTHORS B7-like molecules and uses thereof
 TITLE Patent: WO 0200710-A 3 03-JAN-2002;
 JOURNAL

FEATURES Angen, Inc. (US)
 Location/Qualifiers
 source 1..1168
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 8..1168
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28217.1"
 /translation="MVAGAMENRDPGSGNEVIRGPNARYLKSGAIRENCTYSOG
 WKILMALSPMYLSVRAPEPILITNDPSPSPAPDGGANTSEKLIHNVPSGNTRC
 SLONSRHGSATVLTVOYMGELFTPSNLYVAENPEVTCLEPWTWLPDISBELGL
 VSHSYEVEPESDILQSAVSLIALTPQSNLTICVATWMLA#KSAIVNLVTRCPQ
 DTGGGINIPGLVSLPGLSPPTKRWKVLGAGMLLPTCIITRCCCRRCRCG
 NCCRCFCRRKRGFRIDOKSEKETWKEETESGNSGINSDEKRTDTASLP
 PKSESDPEQRNSCGPPHQRADOPPRASHPOASFNILASPEKSNITTV"
 BASE COUNT 299 a 302 c 288 g 279 t
 ORIGIN
 Query Match 95.5%; Score 1122.4; DB 6; Length 1168;
 Best Local Similarity 99.5%; Pred. No. 9.7e-260;
 Matches 1126; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 44 COTCCAGGCTTCTGGGTCTGTAATGAAGTATGAAGGCCCCAGATGCACAGCTCT 103
 |||||||
 DB 37 CCACACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96
 |||||||
 QY 104 GAAGGCTCCAGGCTGCTCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 163
 |||||||
 DB 97 GAAGGCTCCAGGCTGCTCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 156
 |||||||
 QY 164 GGCCTCAGTACATGAGTGTGCTTAAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 223
 |||||||
 DB 157 GGCCTCAGTACATGAGTGTGCTTAAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 216
 |||||||
 QY 224 CCGCTTCACTCTCAGAGTACGACACGAGCGGAGAACCTTCACTCTGAGATGATCATCA 283
 |||||||
 DB 217 CCGCTTCACTCTCAGAGTACGACACGAGCGGAGAACCTTCACTCTGAGATGATCATCA 276
 |||||||
 QY 284 CAATGTGAGGCCAGTGTGTTGGGGAACATCAGATGAGACCTCCAGAAAGTCCCTGCA 343
 |||||||
 DB 277 CAATGTGAGGCCAGTGTGTTGGGGAACATCAGATGAGACCTCCAGAAAGTCCCTGCA 336
 |||||||
 QY 344 TGGATCTGCTTACCTTACCTGCTCAAGTTATGGAGAGCTGTTCACTCCAGTGTATCT 403
 |||||||
 DB 337 TGGATCTGCTTACCTTACCTGCTCAAGTTATGGAGAGCTGTTCACTCCAGTGTATCT 396
 |||||||
 QY 404 TGTACTGCTGAGATGACCTTGTGAAGTACTTGTCTACCTCAGATCAGATGAGCCGCT 463
 |||||||
 DB 397 TGTACTGCTGAGATGACCTTGTGAAGTACTTGTCTACCTCAGATCAGATGAGCCGCT 456
 |||||||
 QY 464 CCGGATATTTCTCGGAGCTCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
 |||||||
 DB 457 CCGGATATTTCTCGGAGCTCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
 |||||||
 QY 524 GGAGCCAGCAGACCTTCAAAAGTGAAGTGAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCT 583
 |||||||
 DB 517 GGAGCCAGCAGACCTTCAAAAGTGAAGTGAAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
 |||||||
 QY 584 GACTTGTACTTGTGCTGCTTACTGAGAGAGCTTGAAGGCCCGGCAAGTCTGCACTGTAAA 643
 |||||||
 DB 577 GACTTGTACTTGTGCTGCTTACTGAGAGAGCTTGAAGGCCCGGCAAGTCTGCACTGTAAA 636
 |||||||
 QY 644 TCTCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
 |||||||
 DB 637 TCTCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 |||||||
 QY 704 ATCAAGTTTACGAGTTTATGCTTTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763
 |||||||
 DB 697 ATCAAGTTTACGAGTTTATGCTTTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
 |||||||
 QY 764 AGCAGGCCATGCTTCTGACGCCGAGCTGATCTTAACATACGCTGCTGCTGCTGCTGCTGCT 823
 |||||||


```

/clone="PLACE6017788"
/issue_type="placenta"
/clone.lib="PLACE6"
/notes="cloning vector: PME18SFL3"
BASE COUNT      580 a      463 c      474 g      534 t
ORIGIN

```

```

Query Match      95.8% Score 1125.4; DB 9; Length 2051;
Best Local Similarity 99.9%; Pred. No. 1.9e-260;
Matches 1126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 49 ACGGTTGCGGTCTGTAATGAAAGTCAATAGAGGCCCCAGAAATGACAAGTCTCTGAAG 108
DB 201 ACGGTTGCGGTCTGTAATGAAAGTCAATAGAGGCCCCAGAAATGACAAGTCTCTGAAG 260
QY 109 GCTCCAGAGGCTGCTCAACGCAACGCTCCCGGGCTGGAAGCTCATCATGCGGCTC 168
DB 261 GCTCCAGAGGCTGCTCAACGCAACGCTCCCGGGCTGGAAGCTCATCATGCGGCTC 320
QY 169 TCAGTGAATGATGCTGCTAAGCTCAAGGCCCCATGAGCCCATCATCACCATGACCGCT 228
DB 321 TCAGTGAATGATGCTGCTAAGCTCAAGGCCCCATGAGCCCATCATCACCATGACCGCT 380
QY 229 TCACCTCTCAGAGGTACGACGAGGCGGGAACCTTCACCTCGAGATGATCATCCCAATG 288
DB 381 TCACCTCTCAGAGGTACGACGAGGCGGGAACCTTCACCTCGAGATGATCATCCCAATG 440
QY 289 TGGAGCCCAAGTATGCGGGGAACATCAAGTCAAGCCCTCCCAAGACAGTCCGCTGATGAT 348
DB 441 TGGAGCCCAAGTATGCGGGGAACATCAAGTCAAGCCCTCCCAAGACAGTCCGCTGATGAT 500
QY 349 CTGCTTACCTTACCGTCAAGTATGAGAGAGCTGTTTCAATCCAGTGAATCTTGTAG 408
DB 501 CTGCTTACCTTACCGTCAAGTATGAGAGAGCTGTTTCAATCCAGTGAATCTTGTAG 560
QY 409 TCCTGGAATGAACCTTGTGAAGTACTTGTCTACCTTCACACTGGAGCCGCGCTCCGG 468
DB 561 TCCTGGAATGAACCTTGTGAAGTACTTGTCTACCTTCACACTGGAGCCGCGCTCCGG 620
QY 469 ATATTCTCGGAGCTGCTGCTCTCTGTCAGCCATTCACACTTATTTTGTTCGGAGC 528
DB 621 ATATTCTCGGAGCTGCTGCTCTCTGTCAGCCATTCACACTTATTTTGTTCGGAGC 680
QY 529 CCAGCGACCTTCAAGTGAAGTGAAGTCTGCTGCTGAGCCCAAGAGCAATGAGGACT 588
DB 681 CCAGCGACCTTCAAGTGAAGTGAAGTCTGCTGCTGAGCCCAAGAGCAATGAGGACT 740
QY 589 TGACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
DB 741 TGACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
QY 649 CTGCTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
DB 801 CTGCTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
QY 709 GTTACCGAGTTGAGTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
DB 861 GTTACCGAGTTGAGTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
QY 769 GCACCATGCTTCTGACGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
DB 921 GCACCATGCTTCTGACGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
QY 829 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
DB 981 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
QY 889 TTGCTATTCATTTCAAAAGAAATCTGAAAAGAGAAACAAAGAAAGAGAGAGAG 948
DB 1041 TTGCTATTCATTTCAAAAGAAATCTGAAAAGAGAAACAAAGAAAGAGAGAGAG 1100
QY 949 AAAGTGAATGAAGAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008

```

```

DB 1101 AAGGTGAATGAAGAAATCCCGCTACAAATTCAGATGAAACAAAAGACCAACAGACCGCT 1160
QY 1009 CTCTCCCTCCCAATTCCTGTAATCAGTATCTGAAACAAAGAAAGAGTACTGCTGCC 1068
DB 1161 CTCTCCCTCCCAATTCCTGTAATCAGTATCTGAAACAAAGAAAGAGTACTGCTGCC 1220
QY 1069 CTCTCCCTCCCAATTCCTGTAATCAGTATCTGAAACAAAGAAAGAGTACTGCTGCC 1128
DB 1221 CTCTCCCTCCCAATTCCTGTAATCAGTATCTGAAACAAAGAAAGAGTACTGCTGCC 1280
QY 1129 TTAATCTGCGCAGTCTGAGAAAGTCAATTAATCAATCTGATATAG 1175
DB 1281 TTAATCTGCGCAGTCTGAGAAAGTCAATTAATCAATCTGATATAG 1327
RESULT 3
AX380400      1240 bp      DNA      linear      PAT 18-MAR-2002
LOCUS
DEFINITION    Sequence 5 from Patent WO0200710.
ACCESSION     AX380400
VERSION        AX380400.1 GI:19575330
KEYWORDS
SOURCE
ORGANISM      human.
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1             Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
               B7-like molecules and uses thereof
               Patent: WO 0200710-A 5 03-JAN-2002;
               Amgen, Inc. (US)
FEATURES
source        1..1240
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /protein_id="CAD28218.1"
               /note="unnamed protein product"
               /codon_start=1
               /db_xref="GI:19575331"
CDS
               80..1240
               /note="B7-like protein product"
BASE COUNT    318 a      319 c      305 g      298 t
ORIGIN
Query Match      95.6% Score 1123.4; DB 6; Length 1240;
Best Local Similarity 99.9%; Pred. No. 5.6e-260;
Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 51 GCTTCTGCTGCTGTAATGAAGTCAATAGAGGCCCCAGAAATGACAAGTCTCTGAAGGC 110
DB 116 GCTTCTGCTGCTGTAATGAAGTCAATAGAGGCCCCAGAAATGACAAGTCTCTGAAGGC 175
QY 111 TCCAGAGCTGCTTCAATGCTACGCTCTTCCAGGCTGGAAGCTCATCATGCTGCTCTC 170
DB 176 TCCAGAGCTGCTTCAATGCTACGCTCTTCCAGGCTGGAAGCTCATCATGCTGCTCTC 235
QY 171 AGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 230
DB 236 AGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
QY 231 ACCTCTCAGAGGTACGACGAGGCGGGAACCTTCACCTCGAGATGATCATCCCAATG 290
DB 296 ACCTCTCAGAGGTACGACGAGGCGGGAACCTTCACCTCGAGATGATCATCCCAATG 355
QY 291 GAGCCAGTGAATCGGGGAACATCAAGTCAAGCTCCAGAAACAGTCCGCTCATGATGATCT 350
DB 356 GAGCCAGTGAATCGGGGAACATCAAGTCAAGCTCCAGAAACAGTCCGCTCATGATGATCT 415

```

Amgen, Inc. (US)
 Location/Qualifiers
 source 1..1175
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 27..1175
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28216.1"
 /db_xref="GI:19575327"
 /translation="MGLVFLHSGSGNEVIEGPONATVYLGSGARENCYISQCKLI
 MMALSDMYLVSRMPELITNDRTSOPRYDQGNFTSMILHNEPDSOGNIRKSLN
 SRLHGSAYLVQVMELEIPSYNLVAENEDCEVCLPSHMTRLPDISMGLVSHS
 SYFVPEPSDLSAVSIALIPQSNGLTCAVTMSLKARKSAVNLVIRCPDPTGG
 GINIGVLSLESLSPETMGKVLGLAGTLLTPTCTLRCCCGRRCCGNCNC
 RCCCRKRREPRRIQFOKKSEKNTKNETESGNGNSNDEKTDTSALPDKSC
 ESSDEQRNNSCGPPHQRADQRPAPSHAPQAFMLASPEVSNITTV"
 BASE COUNT 295 a 307 c 286 g 287 t
 ORIGIN
 Query Match 100.0%; Score 1175; DB 6; Length 1175;
 Best Local Similarity 100.0%; Pred. No. 2,1e-272;
 Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGTCTGCCATCTGATATAACAAGAGATGAGGCTGTGATTTCTCCACAGGTTCTGGGT 60
 1 CTGTCTGCCATCTGATATAACAAGAGATGAGGCTGTGATTTCTCCACAGGTTCTGGGT 60
 DB 1 CTGTCTGCCATCTGATATAACAAGAGATGAGGCTGTGATTTCTCCACAGGTTCTGGGT 60
 QY 61 CTGTCTGATGAGTATGAG 120
 61 CTGTCTGATGAGTATGAG 120
 DB 61 CTGTCTGATGAGTATGAG 120
 QY 121 GCTTCACTGACAGCGTCTCCAGGCTGGAAGCTCATCATGTGGCTCTCAGTACATGAG 180
 121 GCTTCACTGACAGCGTCTCCAGGCTGGAAGCTCATCATGTGGCTCTCAGTACATGAG 180
 DB 121 GCTTCACTGACAGCGTCTCCAGGCTGGAAGCTCATCATGTGGCTCTCAGTACATGAG 180
 QY 181 TGTGTCTAAGCGTCAAGGCGCCATGAGCCATCATCACCAATGACCGCTTCTACCTCTCA 240
 181 TGTGTCTAAGCGTCAAGGCGCCATGAGCCATCATCACCAATGACCGCTTCTACCTCTCA 240
 DB 181 TGTGTCTAAGCGTCAAGGCGCCATGAGCCATCATCACCAATGACCGCTTCTACCTCTCA 240
 QY 241 GGTTCGACGACGAGGCGGAGCTCACTCGGAGAGATCATCCCAATGATGAGGAGCCAGTG 300
 241 GGTTCGACGACGAGGCGGAGCTCACTCGGAGAGATCATCCCAATGATGAGGAGCCAGTG 300
 DB 241 GGTTCGACGACGAGGCGGAGCTCACTCGGAGAGATCATCCCAATGATGAGGAGCCAGTG 300
 QY 301 ATTGCGGGAACATCATGATGACAGCTCCAGAACAGTGCCTCATGATGATGATGATGAT 360
 301 ATTGCGGGAACATCATGATGACAGCTCCAGAACAGTGCCTCATGATGATGATGATGATGAT 360
 DB 301 ATTGCGGGAACATCATGATGACAGCTCCAGAACAGTGCCTCATGATGATGATGATGATGAT 360
 QY 361 CCGTCCAAAGTTATGAG 420
 361 CCGTCCAAAGTTATGAG 420
 DB 361 CCGTCCAAAGTTATGAG 420
 QY 421 AACCTTGGAAGTACTTGTCTACCTCCACACAGGAGAGAGAGAGAGAGAGAGAGAGAG 480
 421 AACCTTGGAAGTACTTGTCTACCTCCACACAGGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 421 AACCTTGGAAGTACTTGTCTACCTCCACACAGGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 AGCTCGGCTCTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAGAGAG 540
 481 AGCTCGGCTCTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAGAGAG 540
 DB 481 AGCTCGGCTCTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAGAGAG 540
 QY 541 AAATGTCAGTATGAGTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAG 600
 541 AAATGTCAGTATGAGTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAG 600
 DB 541 AAATGTCAGTATGAGTCTGTCAGCAGCATTAAGTATTTTGTTCGAGAGAGAGAGAG 600
 QY 601 CTACCTGGAAGAGCTGTAAG 660
 601 CTACCTGGAAGAGCTGTAAG 660
 DB 601 CTACCTGGAAGAGCTGTAAG 660
 QY 661 GTCCCAAGAGACATGAG 720
 661 GTCCCAAGAGACATGAG 720
 DB 661 GTCCCAAGAGACATGAG 720
 QY 721 TAGGTTTTTCATGCTACTTGGGGCAAGTTGAGACTTGAGAGAGAGAGAGAGAGAGAG 780
 721 TAGGTTTTTCATGCTACTTGGGGCAAGTTGAGACTTGAGAGAGAGAGAGAGAGAGAG 780

DB 721 TAGGTTTTTCATGCTACTTGGGGCAAGTTGAGACTTGAGAGAGAGAGAGAGAGAGAG 780
 QY 781 TGAGCGGAGAGTACTTACTTACATATACATACATACATACATACATACATACATACATAC 840
 DB 781 TGAGCGGAGAGTACTTACTTACATATACATACATACATACATACATACATACATACAT 840
 QY 841 GCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 841 GCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 901 TTCAAAAGAAATCTGAAAAG 960
 DB 901 TTCAAAAGAAATCTGAAAAG 960
 QY 961 AAAACTCCGGCTAATTAATTCAGATGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 DB 961 AAAACTCCGGCTAATTAATTCAGATGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1021 AATCCTGTGAATTCAGATGATCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 DB 1021 AATCCTGTGAATTCAGATGATCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 QY 1081 GGGCTGATCAACGTCACCCAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 DB 1081 GGGCTGATCAACGTCACCCAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 QY 1141 GTCTGAGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1175
 DB 1141 GTCTGAGAGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1175

RESULT 2
 AK092516 2051 bp mRNA linear PRI 15-JUL-2002
 LOCUS AK092516
 DEFINITION Homo sapiens cDNA FLJ35197 fis, clone PLACE6017788, highly similar
 to IGS75.
 ACCESSION AK092516
 VERSION AK092516.1 GI:21751130
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens placenta cDNA to mRNA, clone_11b:PLACE6
 clone:PLACE6017788.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagaitsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
 Oshima, A., Suzuki, Y., Sugano, S., Nagaharai, K., Masuno, Y., Nagai, K.
 and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 TITLE NEDO human cDNA sequencing project
 JOURNAL 2 (bases 1 to 2051)
 REFERENCE Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA library
 Research Association for Biotechnology (RAB); (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing; HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source 1..2051
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 23:25:03 ; Search time 3244.94 Seconds

(without alignments) 10538.186 Million cell updates/sec

Title: US-09-729-264-1

Perfect score: 1175

Sequence: 1 cctctgcacatcgtatataa.....gtaatacaactgtatagatag 1175

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_of:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rdt:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hggo_hum:*
40: em_hggo_mus:*
41: em_hggo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1175	100.0	1175	6	AX380396	AX380396 Sequence
2	1125.4	95.8	2051	6	AK092516	AK092516 Homo sapi
3	1123.4	95.6	1240	6	AX380400	AX380400 Sequence
4	1122.4	95.5	1168	6	AX380398	AX380398 Sequence
5	1067.4	90.8	1139	6	AX380402	AX380402 Sequence
6	381	32.4	1195	6	AX380404	AX380404 Sequence
7	320.2	27.3	142742	9	AF121782	AF121782 Homo sapi
8	320.2	27.3	340000	9	HS212080	HS212080 Homo sapi
9	188.6	16.1	196900	2	AC020851	AC020851 Mus muscu
10	186	15.8	754	6	AX380408	AX380408 Sequence
11	186	15.8	895	6	AX380406	AX380406 Sequence
12	186	15.8	1556	10	BC004806	BC004806 Mus muscu
13	177.6	15.1	175861	2	AC120346	AC120346 Mus muscu
14	100	8.5	199665	9	AF064857	AF064857 Homo sapi
15	58	4.9	175861	2	AC120346	AC120346 Mus muscu
16	58	4.9	196900	2	AC020851	AC020851 Mus muscu
17	57	4.9	40205	9	AF045450	AF045450 Homo sapi
18	57	4.9	170121	9	AF064860	AF064860 Homo sapi
19	56.6	4.8	427	9	HS063332	HS063332 Human super
20	55.4	4.7	53908	2	AC120145	AC120145 Mus muscu
21	55	4.7	1061	9	HSCAGCTG	AC100501 Homo sapi
22	53	4.5	178052	2	AC118358	AC118358 Rattus no
23	52.8	4.5	138685	2	AC017582	AC017582 Drosophi
24	52.8	4.5	166249	3	AC093502	AC093502 Drosophi
25	52.8	4.5	185404	3	AC104703	AC104703 Drosophi
26	52.8	4.5	341319	3	AE003536	AE003536 Drosophi
27	51.8	4.4	397	6	BD010479	BD010479 Diagnost
28	51.4	4.4	85584	2	AC099238	AC099238 Rattus no
29	50.8	4.3	162971	2	AC103439	AC103439 Rattus no
30	50.6	4.3	147185	9	AC020589	AC020589 Homo sapi
31	50.6	4.3	165675	2	AC113925	AC113925 Rattus no
32	50.6	4.3	195460	2	AC128068	AC128068 Rattus no
33	50	4.3	94674	2	AC128466	AC128466 Rattus no
34	50	4.3	179494	2	AC113770	AC113770 Rattus no
35	50	4.3	185487	2	AC108346	AC108346 Rattus no
36	49.8	4.2	299	9	HS131896	HS131896 Homo sapi
37	49.8	4.2	161170	2	AC129166	AC129166 Rattus no
38	49.4	4.2	49845	2	AC100974	AC100974 Mus muscu
39	49.4	4.2	160650	3	AC008188	AC008188 Drosophi
40	49.4	4.2	160885	2	AC018184	AC018184 Drosophi
41	49.4	4.2	162384	2	AC007840	AC007840 Drosophi
42	49.4	4.2	172334	2	AC122602	AC122602 Rattus no
43	49.4	4.2	173373	3	AC099016	AC099016 Drosophi
44	49.4	4.2	306854	3	AE003797	AE003797 Drosophi
45	49.4	4.2	312634	2	AC095096	AC095096 Rattus no

ALIGNMENTS

RESULT 1
AX380396
LOCUS AX380396 1175 bp DNA
DEFINITION Sequence 1 from Patent WO0200710.
ACCESSION AX380396
VERSION AX380396.1 GI:19575326
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Welcher A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE BV-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 1 03-JAN-2002;

[illegible]

Result	ID	Standard	DNA	BP
RESULT 13	AAx89891/c	AAx89891 standard	DNA	397 BP
AC	AAx89891;			
DT	05-NOV-1999	(first entry)		
DE	Spinocerebellar ataxia type III (SCAIII) gene fragment.			
KM	Spinocerebellar ataxia type III; SCAIII; reverse dot hybridisation;			
KM	PCR-microplate hybridisation; PCR-MPH; trinucleotide repeat; TNR;			
KM	SCAIII syndrome; ss.			
OS	Homo sapiens.			
Key	Location/Qualifiers			
FT	repeat_region	137..355		
FT	/*tag=	a		
FT	repeat_unit	137..139		
FT	/*tag=	b		
FT	/note=	"trinucleotide repeat"		
PN	W09943852-A1.			
PD	02-SEP-1999.			
PF	18-FEB-1999;	99WO-KR00078.		
PR	26-FEB-1998;	98KR-0006278.		
PA	(JIND/) JIN D K.			
PA	(SMSU) SAMSUNG FINE CHEM CO LTD.			
PI	Jin DK;			
DR	WPI; 1999-527634/44.			
PT	Diagnosis of spinocerebellar ataxia type III (SCA III) syndrome			
PT	using techniques which ensure highly accurate diagnosis			

[illegible]

ID	ABK10240 standard; DNA; 403 BP.
XX	ABK10240;
XX	20-MAY-2002 (first entry)
XX	Trinucleotide repeat sequence #3.
XX	Trinucleotide repeat; fragile X syndrome; ds;
XX	spino cerebellar ataxia type II; DRPLA; neuro-psychiatric;
XX	hereditary hypertrophic cardiomyopathy; Marfan syndrome;
XX	dentatorubral and pallidolysian atrophy;
XX	spino cerebellar ataxia; X-linked spinobulbar atrophy.
OS	Homo sapiens.
XX	Key
XX	Location/Qualifiers
XX	repeat_region
XX	152..347
XX	/*tag=a
XX	/rpt_type=TANDEM
XX	152..154
XX	/*tag=b
XX	/note="CAG type repeat"
OS	Homo sapiens.
XX	RR2000003004-A.
XX	15-JAN-2000.
XX	25-JUN-1998; 98KR-0024064.
XX	25-JUN-1998; 98KR-0024064.
XX	(SMSU) SAMSUNG MEDICAL CENT.
XX	(JIND/) JIN D G.
XX	Jin DG;
XX	

RESULT 8
 ID ABK13034 standard; cDNA: 754 BP.
 AC ABK13034;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding mouse B7-1-like protein, B7-L_{m3}.
 KM Mouse; B7-1-like protein; B7-L; antifertility; gynaecological;
 KM antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KM antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;
 KM antischistosomal; nephrotoxic; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; neuropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 53..724
 FT /*tag= a
 FT /product= "B7-1-like protein, B7-L_{m3}"
 XX
 XX MO200200710-A2.
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 DR WPI: 2002-130881/17.
 DR P-PSDB; AAU75546.
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 1; Fig 7; 135pp; English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ

CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or all sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of mouse B7-L_{m3}.
 XX
 SO Sequence 754 BP; 220 A; 191 C; 175 G; 168 T; 0 other;
 Query Match 15.8%; Score 186; DB 24; Length 754;
 Best Local Similarity 71.1%; Pred. No. 3.2e-41;
 Matches 246; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 Db 24 GAGATGGGCTTGATTTCTCCACGGTCTGGCTGTPAATGATACAGAGGC 83
 Db 83 GTGCTGTATCTCTGCTCAGCTGACAGCTCCGGATCCAGTATACATACAGAGGT 142
 QY 84 CCCAGATGCAACAGTCTGAAAGGCTCCAGCTCGTTCACTGACCGCTTCCAG 143
 Db 143 CTCAGATGTATACAGTCTAAGAGCTAGAGGCTCAGTCACTCACTGACCGTCACTCAC 202
 QY 144 GGTGGAGAGCTCATCTGTCGCTCAGAGATAGTGTGTAGCGTCAAGGCCATG 203
 Db 203 GGTGGAGAGCTCATCTGTCGCTCAGAGATAGTGTGTGTAGCGTCAAGGCCATG 262
 QY 204 GAGCCCATCATCACAATGACCGCTTCACTCAGAGGATGACCGAGGCGGAGACTTC 263
 Db 263 GAGCCCATCATCACAATGACCGCTTCACTCAGAGGATGACCGAGGCGGAGACTTC 322
 QY 264 ACCTGGAGATGATCATCCAAATGTGGAGCCGATGATGGGGGAATCATGATGACG 323
 Db 323 ATCTGGAGATGATCATCCAAATGTGGAGCCGATGATGGGGGAATCATGATGACG 382
 QY 324 CTCAGAGACAGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATG 369
 Db 383 CTCAGAGACAGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATG 428
 RESULT 9
 ID ABK13033 standard; cDNA: 895 BP.
 AC ABK13033;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding mouse B7-1-like protein, B7-L_{m2}.
 KM Mouse; B7-1-like protein; B7-L; antifertility; gynaecological;
 KM antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
 KM antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antithyroid; antiulcer; antiallergic;
 KM antischistosomal; nephrotoxic; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; neuropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 53..865
 FT /*tag= a
 FT /product= "B7-1-like protein, B7-L_{m2}"
 XX
 PN MO200200710-A2.
 XX
 PD 03-JAN-2002.

OY	84	CCCCGAATGCAACAGCGCTGGAAAGGGCTCCAGAGCGCTGCATTCACTGCACCGTCTCCAG	143
Dd	143	CCTGAGATGTAAACAGTGCTCTTAAGAAGACTAGAGGGCTCACCTTAACTGACCGCTGACTC	202
OY	144	GCGTGGAGGCTCATCATGTGGGGCTCTCAGTGTGACATGTGTGTCTTAAGCGTCAAGGCCATG	203
Dd	203	GGCTGGAGAGCTTCATGTGAGACTCTTAACAATAATGTGTGTGTGCTGAGTCTCACCAACCCA	262
OY	204	GAGCCCATCATCAACCAANTGACCGCTTCACCTCTCAGAGGTACAGACCAGGCGGGAACTTC	263
Dd	263	GGAGCCCATCATCAACCAACCGCTTCACCTCTGAGTTATCAACAGCACTGACAGAGCTTC	322
OY	264	ACCTCGAGATGTGATCATCCCATGTGGAGCCAGAGTATGGGGAAATCATGATGATGACG	323
Dd	323	ATCTCGAGATGTGATCATCCCATGTGTGACAGCCCATGACTGCGGATCCGCTCAATTGCAAC	382
OY	324	CTCCAGAAGAGTGCCTGCCTCATGTGATCTGCTTACCTTACCGTCAAATTATGGGAGAGCTG	383
Dd	383	CTCGAAGAACAGCCATGGGTTTGGATGTGGCTTCCTCTCAGTCAAGTATCATGGGAGCCCTG	442
OY	384	TTCATTCCCAGTGTATCTGTGTAGTCGTGTGAANTGAACCTGTGAGATTACTGTCTA	443
Dd	443	AACATTCCTTGACAAACAACCTTATGTACTGAGGGGAACCTGTAAATGTACTGTCTAT	502
OY	444	CCCTCACACTGGAACCCGGCTCCCGGATATTTCCTGGGAGCTTGCTCTCCTGGTAGCCAT	503
Dd	503	GGCTGGGCTGGAGACTCATCTCCCGGATATTTCCTGGAGACTTGAGGTTCCCGTAAGCAT	562
OY	504	TCAAGCTATTATTTTTGTTCGGGAGCCGACAGACTTCAAAGTGTGACATCTGTGCT	563
Dd	563	TCGAGTTACATTTCTTTGTGGAGCCGGGCAACTTATGAGGGTCTTAAGTGTCTTGAGAC	622
OY	564	CTGACCCCAAGAGACANTGTGGACTTTGACTTGGCTGGTCTACTGTGAAGAAGCTGTGAAGCC	623
Dd	623	CTCACACCACTGGGCAACGGGAGCCTTGACTGTGTGTGGAGACTBAAGACCTGTGAGGCC	682
OY	624	CGCAAGTCTGCACCTGTAAATCTCACTGTGATTCGTTGGTGTCCCAAGACACTGTGAGGTGT	683
Dd	683	AGCAAGTCTTAACTGTCAACCTGTGAGTGTGTACCCCTCCACCTGACAGATGTGGAGAG	742
OY	684	ATTAAATTCAGAGTGTATTTATCAAGATTACGAGTTTGAAGTTTTCATGTGGCTACTTGG	743
Dd	743	GA-----AGGCCCAAGCACTGCCACCTGG	766
OY	744	GGCAAAAGTTGGACTTGGACTGACAGGACACCATCTTCTGACCGCAGAGCTGTACTTTACA	803
Dd	767	GGCATCATCTGCTGGCAGTGGGCTTCTCTCTCTCTGATCTGTATCATTTGTTTGAT	826
OY	804	ATTACGCTCTGCTCTGCCCGCTCGTTGTGTGGGTGCACTGCTGCTGGCGTTGTGT	863
Dd	827	ATAATATTCTGTCTCTGTGT-----GCTCTCAGAGAGAGAAAAGGAAGATC	873
OY	864	TTTCGCTGTAGAAAAAAGAGATTTTCGATTCAATTTCAAAAAAATCTGAAAAAGAG	923
Dd	874	TACTTATCAAAATGAATTAAGGAATCTGCAAAACATGAGSACA-----AC	919
OY	924	AAGACAAACAAAGAACTGAGACAGAAAGTGGAAATGAAAATCCGGCTACATTTCAAT	983
Dd	930	AAAGCAGATCCGGGAGACAAAGTTAAAAAGTGAAGAAAGAAACATACGGGTACAGTTCGAT	979
OY	984	GAACAAAGACACAGACACCGCTGCTGCTCCGCCAAATCCGTGAATCCAGTATGATCT	1043
Dd	980	GAGGCAAAAGGCTGACAGACTGCTCTCTCTCTCTAAATCTGTCTGAAGTCAACGCTTCCA	10399
OY	1044	GAACAAAGAAACAGTACGCTGTGGCCCTCTCACACAGCGGGCTGATCAACGTCACCCAGG	11033
Dd	1040	GA AAAACG CAGCAGTA-----GCCTTCCTTACAGAGAACTCAATAAACAATCAGCCCGT	10933
OY	1104	CGAGCAATCATCCACAGGCTCTTTTAATCTGCCAAGTCTGAGAAAGSTGATGTAATACA	11533
Dd	1094	CCAGCAACTCATCAAGGGTTTCTCTTGAATCGCCAGTCTCTCAGAAAGTCTGAAATGTG	11533

QY	1164	ACTGTAGTATA	1174
Db	1154	ACTTTAGTGTA	1164

RESULT 7
AAI36582
ID AAI36582 standard; DNA; 401 BP.

AC AAI36582;

DT 17-OCT-2001 (first entry)

DE Probe #5268 used to measure gene expression in human placenta sample.

KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.

XX
XX
OS Homo sapiens.

AA W0200157272-A2.
PN

AA PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US006663.

PR 04-FEB-2000; 2000US-0180312.

PR 30-JUN-2000; 2000US-0608408.

PR 21-SEP-2000; 2000US-0234687.
PB 27-SEP-2000; 2000US-0236359

PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

XX
P1
Fenn 39, hanzel da, chich w, nann da

XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

PT analyzing gene expression in human placenta -

PS Claim 25; SEQ ID No 5268; 654pp; English.

CC The present invention relates to single exon nucleic acid probes (SENP).

CC producing a microarray for predicting, measuring and displaying gene
CC expression is a sample derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.
XY

sequence 401 BP; 105 A; 98 C; 111 G; 8/ T; 0 other,

Query Match	Score	DB sz	Length
Best Local Similarity	97.08;	Pred. No. 1.4e-42;	

[illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

331

163 mcc 221

Dh
393 TTTTCTACATTTCATGAGCCACTCATCAATAAAT 381

04 222 GAGCGCTTCACTTCTAGAG 241

Db 382 GACCGGTTCACTCTCAGAG 401

Db	301	CGGCTCCGGATATTTCCTGGAGACTGGCTCTCCTGGTCAGCATTTCAAGCTATTATTTT		360
OY	519	GTTCGGAGACCACGACCTTCAAAAGTCAGTAGCATCCCTGGCTCTGACCCACAGAGC		578
Db	361	GTTCGGGGGCCACGACCTTCAAAAGTCAGTAGCATCCTGGCTCTGACCCACAGAGC		420
OY	579	AATGGAGCTTTCAGTTCCGGGGCTACCTGGAAAGCTCAAGGCCCGCAAGTCTCACT		638
Db	421	AATGGAGCTTTCAGTTCCGGGGCTACCTGGAAAGCTCAAGGCCCGCAAGTCTCACT		480
OY	639	GTAATCTCACTGTGATTCCGGTGTCCCAAGACATCGAGGTGTATTAATATTCAGGT		698
Db	481	GTAATCTCACTGTGATTCCGGTGTCCCAAGACATCGAGGTGTATTAATATTCAGGT		540
OY	699	GTAATATCAAGTTACCGAGTTTAGTTTTTCAATGCCACTTTGGGGCAAAATTGGACTT		758
Db	541	GTAATATCAAGTTACCGAGTTTAGTTTTTCAATGCCACTTTGGGGCAAAATTGGACTT		600
OY	759	GGACTACAGGACACATGCTTTCGACGCCAGCTGTACTTACAATAGCTGCTGTGC		818
Db	601	GGACTACAGGACACATGCTTTCGACGCCAGCTGTACTTACAATAGCTGCTGTGC		660
OY	819	TGCGCCGCTGTGTGTGTGGCTGCACATGCTGTGCGCTGTGTGTTTGTGCTGTAGAAGA		878
Db	661	TGCGCCGCTGTGTGTGTGGCTGCACATGCTGTGCGCTGTGTGTTTGTGCTGTAGAAGA		720
OY	879	AAAAGAGATTTGCTATTC-----		897
Db	721	AAAAGAGGTGACATTTCTTGTGACATCTTTCCACATCCCAAGATTAGTGCATGATC		780
OY	898	-----	-----	897
Db	781	AATGGCGTCAATTTGTGTATACATATTATGTGTGCTCTTGAGGCTCTCACTGTAGA		840
OY	898	-----AATTTCAAAAGAAATCTGA AAAAGAGAGAAACA AAAGAACTGAGACAG		948
Db	841	CTGATCCATACCGTGGGCACGAAATCTGA AAAAGAGAGAAACA AAAGAACTGAGACAG		900
OY	949	AAAGTGAATGA AAAAATCTCGGGCTACAAATTCAGATTAACAAAAGACACAGACCGGTT		1008
Db	901	AAAGTGAATGA AAAAATCTCGGGCTACAAATTCAGATTAACAAAAGACACAGACCGGTT		960
OY	1009	CTCTCCCTCCCAAAATCTGTGAATTCAGAGTATCTGAACAAAAGAAACAAGTACTGTGCC		1068
Db	961	CTCTCCCTCCCAAAATCTGTGAATTCAGAGTATCTGAACAAAAGAAACAAGTACTGTGCC		1020
OY	1069	CTCTCCACACAGCGGCTGATCAACGT		1094
Db	1021	CTCTCCACACAGAGACATAGACCT		1046

RESULT 6	
ABK13032	
ID	ABK13032 standard; cDNA; 1195 BP.
XX	
AC	ABK13032;
XX	
DT	23-APR-2002 (first entry)
XX	
DE	DNA encoding mouse B7-1-like protein, B7-L_m1.
XX	
XX	Mouse; B7-like protein; B7-L; antiinfertility; gynaecological; antiinflamm; cytostatic; immunosuppressive; antiarthritic; antirheumatic antiinflammatory; dermatological; antipsoriatic; neuroprotective; antiidiabetic; haemostatic; antithyroid; antitumor; antiallergic; reproductive disorder; nephropathy; virulence; tumor; cancer; antiproliferative disorder; graft versus host disease; autoimmune disease; toxic shock syndrome; allergy; nephropathy; skin disorder; endocrinopathy; lymphoproliferative disorder; gene; ss.
XX	
OS	Mus musculus.
XX	

FH	Key	Location/Qualifiers
FT	CDS	53..1165
ET		/tag= a
XX		/product= "B7-like protein, B7-L.m1"
PN		
XX		
XX		
PR		MO200200710-A2.
PF		
XX		
XX		28-JUN-2001; 2001WO-US20719.
XX		
XX		
XX		
XX		
PI		Weicher AA, Sarmiento UM, Schultz HJ, Chute HT;
DR		WPI; 2002-130881/17.
XX		P-PSTDB; AAW05544.
PT		
PT		New B7-like polypeptides, polynucleotides and their modulators, useful
PS		for diagnosing, preventing and treating reproductive, immune and
XX		proliferative disorders, e.g. cancer and arteriosclerosis
XX		
XX		Claim 1; Fig 5; 135pp; English.
CC		The invention relates to an isolated B7-like (B7-L) polypeptide (I).
CC		The polypeptide, polynucleotide encoding it and antibody against (I) are
CC		useful for treating B7-like polypeptide-related disease, disorders or
CC		conditions including reproductive disorders (e.g. Infertility,
CC		miscarriage, preterm labour and delivery and endometriosis) and
CC		proliferative disorders. Antibodies, soluble proteins comprising
CC		extracellular domains and other regulators of B7-L polypeptides are
CC		useful for enhancing the immune response to tumours. (I) plays a role in
CC		growth and maintenance of cancer cells based on the observation of
CC		seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
CC		polypeptide. Hence modulators of (I) are useful for the treatment of
CC		cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
CC		testicular cancer and cancers of hematopoietic system. B7-L polypeptide
CC		pathway can be manipulated to regulate cytotoxic T-lymphocyte response
CC		in allograft transplantation, graft versus host disease, T-cell
CC		dependent B-cell mediated diseases and autoimmune diseases. B7-L
CC		molecules are useful for alleviating the symptoms associated with
CC		diseases involving chronic immune cell dysfunction or to treat
CC		autoimmune diseases such as systemic lupus erythematosus, Rheumatoid
CC		arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
CC		purpura and psoriasis, chronic inflammatory disease such as
CC		inflammatory bowel disease (Crohn's disease and ulcerative colitis),
CC		Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
CC		are also useful as immunosuppressive agents for bone marrow and organ
CC		transplantation or to prolong graft survival. B7-L molecules are also
CC		useful for diagnosis and treatment of diseases involving abnormal cell
CC		proliferation, including arteriosclerosis and vascular restenosis.
CC		Antagonists of B7-L polypeptides are useful for alleviation of toxic
CC		shock syndrome or allo sensitisation due to blood transfusions, and for
CC		treatment of allergy, asthma and hypersensitivity reactions,
CC		nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
CC		pemphigoid), endonophopathies (Grave's disease), various pneumopathes
CC		(extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
CC		thrombocytopoenias, Guillain-Barre syndrome and myasthenia gravis, and
CC		lymphoproliferative disorders such as multiple myeloma. The present
CC		sequence represents the coding sequence of mouse B7-L.m1.
XX		
SQ		Sequence 1195 BP; 313 A; 304 C; 291 G; 287 T; 0 other;
Oetry Match	32.4%;	Score 381; DB 24; Length 1195;
Best Local Similarity	61.9%;	Pred. No. 3.4e-95;
Matches 712; Conservative	0;	Mismatches 370; Indels 69; Gaps
24	GAGATGGGCGCTTGATTTTCCTCCACGGTCTGTGGTGAATGAAGCATAGAAGC	83
I		
83	GTGTGGTCAVCTGGCTCAGCTGCAGCTTCGGCATCTATTTCATGATCAATGAAGCT	142
I		

Db 270 CAATGTGAGACCCAGTATCGGGGAATCATCAGATCCCTCCAGACATGCGCTGCA 329
 QY 344 TGGATCTGCTTACCTTACCGTCCAAAGTATGGAGAGCTGTTTCATTCCTCCAGTTAATCT 403
 Db 330 TGGATCTGCTTACCTTACCGTCCAAAGTATGGAGAGCTGTTTCATTCCTCCAGTTAATCT 389
 QY 404 TGTATGCTGCTGAGATGAACCTTGTGAAGTACTTGTCTACCTCTACACTGAGACCGGCT 463
 Db 390 TGTATGCTGCTGAGATGAACCTTGTGAAGTACTTGTCTACCTCTACACTGAGACCGGCT 449
 QY 464 CCGGGAATTTCTGGAGGTGGTCTCCGTCAGCCACCTCAATCAAGTATATTTTGTCC 523
 Db 450 CCGGGAATTTCTGGAGGTGGTCTCCGTCAGCCACCTCAATCAAGTATATTTTGTCC 509
 QY 524 GGAGCCCGAGGACCTTCAAAAGTGCAGTGCATCTGCGCTGACCCCAAGCAAGCAATGG 583
 Db 510 GGAGCCCGAGGACCTTCAAAAGTGCAGTGCATCTGCGCTGACCCCAAGCAAGCAATGG 569
 QY 584 GACTTGTACTTGGCTGCTACTGAGAGAGCTGAAAGCCCGCAAGTCTGCAACTGTATA 643
 Db 570 GACTTGTACTTGGCTGCTACTGAGAGAGCTGAAAGCCCGCAAGTCTGCAACTGTATA 629
 QY 644 TCTCAGCTGATTCGGTGTCCCAAGCACTGAGGTGATTAATATTCAGGTGTATT 703
 Db 630 TCTCAGCTGATTCGGTGTCCCAAGCACTGAGGTGATTAATATTCAGGTGTATT 689
 QY 704 ATCAAGTTACCGAGTTAGTTTTCATTTGCTACTTGGGCAAGTTGCACTTGACT 763
 Db 690 ATCAAGTTACCGAGTTAGTTTTCATTTGCTACTTGGGCAAGTTGCACTTGACT 749
 QY 764 AGAAGCACCATGCTTCTGAGCGCGAGCTGCTTCTTACAAATACGCTGCTGCTGCCG 823
 Db 750 AGAAGCACCATGCTTCTGAGCGCGAGCTGCTTCTTACAAATACGCTGCTGCTGCCG 809
 QY 824 CCGGCTGTTGTGGGCAACTGCTGCTGCTGCTGTTGTTTGTGTGTAAGAAAGAAAG 883
 Db 810 CCGGCTGTTGTGGGCAACTGCTGCTGCTGCTGTTGTTTGTGTGTAAGAAAGAAAG 869
 QY 884 AGGATTTCTGATTCATTTCAAAAGAAATCTGAAAAAGAGAGACCAAGCAAAAGTGA 943
 Db 870 AG-----GAAATCTGAAAAAGAGAAACCAAAACAAAGCAAACTGA 907
 QY 944 GACAGAAAGTGAATGAAGAAATCTCGGCTTACATTCAGATGAACCAAAAGCCACACAC 1003
 Db 908 GACAGAAAGTGAATGAAGAAATCTCGGCTTACATTCAGATGAACCAAAAGCCACACAC 967
 QY 1004 CGGTTCTCCCTCCCAATCTGTGAATCAGTATCCTGAACCAAGAAAGAGTACGCTG 1063
 Db 968 CGGTTCTCCCTCCCAATCTGTGAATCAGTATCCTGAACCAAGAAAGAGTACGCTG 1027
 QY 1064 TGGCCCTCTCACCAGCGGCTGATCAACGTCACCCAGCCAGCAAGTATCCTCACAGGC 1123
 Db 1028 TGGCCCTCTCACCAGCGGCTGATCAACGTCACCCAGCCAGCAAGTATCCTCACAGGC 1087
 QY 1124 TTCTTTTAACTGGCCAGTCTGAGAAAGTCAATATACACTGTGTATTA 1174
 Db 1088 TTCTTTTAACTGGCCAGTCTGAGAAAGTCAATATACAACTGTGTATTA 1138

RESULT 5

AAS92356 standard; cDNA; 1392 BP.
 ID AAS92356 standard; cDNA; 1392 BP.
 AC AAS92356;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #28160.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.

XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PE 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSEQ-) HYSEQ INC.
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 XX
 P-PSDB; ABG28169.
 PT
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID NO 28160; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1392 BP; 338 A; 361 C; 341 G; 352 T; 0 other;
 Query Match 66.3%; Score 779.2; DB 23; Length 1392;
 Best Local Similarity 87.3%; Pred. No. 1.3e-205;
 Matches 913; Conservative 0; Mismatches 23; Indels 110; Gaps 1;
 QY 159 ATGTGGGCTCTCAGTGAATGATGCTGTGCTAAGCGTCAAGCCCATGAGCCCATCATCACC 218
 Db 1 ATGTGGGCTCTCAGTGAATGATGCTGTGCTAAGCGTCAAGCCCATGAGCCCATCATCACC 60
 QY 219 AATGACCGCTTCACTCTCAGAGTAGACACAGGCGGGAACCTTACCTCGGAATGATC 278
 Db 61 AATGACCGCTTCACTCTCAGAGTAGACACAGGCGGGAACCTTACCTCGGAATGATC 120
 QY 279 ATCCACATGTGAGACCCAGGAGTATGCGGGGAACATCAGATCAGAGCTCCACAAAGTCGC 338
 Db 121 ATCCACATGTGAGACCCAGGAGTATGCGGGGAACATCAGATCAGAGCTCCACAAAGTCGC 180
 QY 339 CTGATGATCTGCTTACCTTACCTCCAGTTATGGAGAGCTGTTCATTCACAGTGT 398
 Db 181 CTGATGATCTGCTTACCTTACCTCCAGTTATGGAGAGCTGTTCATTCACAGTGT 240
 QY 399 AATCTTGTAGTCTGAGAAATGACCTTGTGAAGTACTTGTCTACCTCAGCTGAGCC 458
 Db 241 AATCTTGTAGTCTGAGAAATGACCTTGTGAAGTACTTGTCTACCTCAGCTGAGCC 300
 QY 459 CGGCTCCGGAATATTTCTCGGAGCTCGTCTCCTCGTACAGCAATCAAGTATATTTT 518

QY 644 TCTCATGTGATGCGTGTCCCAAGACACTGGAGTGTATTAATATCCAGGTATT 703
 DB 637 TCTCATGTGATGCGTGTCCCAAGACACTGGAGTGTATTAATATCCAGGTATT 636
 QY 704 ATCAAGTTTACCGAGTTTGGATTTCATTCCTACTTGGGGCAAAAGTTGGACTTGGACT 763
 DB 697 ATCAAGTTTACCGAGTTTGGATTTCATTCCTACTTGGGGCAAAAGTTGGACTTGGACT 756
 QY 764 AGCAGGACCAATGCTTGTGACGGCAGAGCTGCTTACAAATACGCTGCTGCTGCGC 823
 DB 757 AGCAGGACCAATGCTTGTGACGGCAGAGCTGCTTACAAATACGCTGCTGCTGCGC 816
 QY 824 CCGTCTGTTGTTGGCTGCAACTGCTGCTGCTGTTGTTGCTGCTGAGAAAGAAAG 883
 DB 817 CCGTCTGTTGTTGGCTGCAACTGCTGCTGCTGTTGTTGCTGCTGAGAAAGAAAG 876
 QY 884 AGCATTTGTAATTCATTTCAAAAAGAAATCTGAAAAAGACAAACAAAGAAACTGA 943
 DB 877 AGCATTTGTAATTCATTTCAAAAAGAAATCTGAAAAAGACAAACAAAGAAACTGA 936
 QY 944 GACAGAAAGTGAATGAATAATCCGCTCAATTCAGATGAACAAAGCCACAGACAC 1003
 DB 937 GACAGAAAGTGAATGAATAATCCGCTCAATTCAGATGAACAAAGCCACAGACAC 996
 QY 1004 CGCTTCTCTCCCTCCCAAAATCTGTGAAATCCAGTATCCTGAAACAAAGAAAGAAAGCTG 1063
 DB 997 CGCTTCTCTCCCTCCCAAAATCTGTGAAATCCAGTATCCTGAAACAAAGAAAGAAAGCTG 1056
 QY 1064 TGGCCCTCTCCACAGCGGGCTGATCAAGTCCACCCAGCCAGCAAGTATCCACAGCG 1123
 DB 1057 TGGCCCTCTCCACAGCGGGCTGATCAAGTCCACCCAGCCAGCAAGTATCCACAGCG 1116
 QY 1124 TTCTTTTAATCTGGCCAGTCTGTGAAAGTCAATTAATCAACATGTTGATATG 1175
 DB 1117 TTCTTTTAATCTGGCCAGTCTGTGAAAGTCAATTAATCAACATGTTGATATG 1168

RESULT 4
 ABK13031
 ID ABK13031 standard; cDNA: 1139 BP.
 AC ABK13031;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding human B7-1-like protein, B7-L_{h4}.
 XX
 KW Human; B7-1-like protein; B7-L; antiinfectivity; gynaecological;
 KW antitumour; cytostatic; immunosuppressive; antiarthritic; antineoplastic;
 KW antiinflammatory; dermatologic; antiproliferative; neuroprotective;
 KW antidiabetic; haemostatic; antihypertensive; antileukemic; antiallergic;
 KW antidiabetic; nephropathic; antibacterial; virucide; tumour; cancer;
 KW reproductive disorder; graft versus host disease; autoimmune disease;
 KW toxic shock syndrome; allergy; nephropathy; skin disorder;
 KW endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1134
 FT /tag a
 FT /product "B7-1-like protein, B7-L_{h4}"
 XX
 PN WO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX

PA (AMGE-) AMGEN INC.
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 XX WPI; 2002-130881/17.
 DR
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 1; Fig 4; 13pp; English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (1).
 CC The polypeptide, polynucleotide encoding it and antibody against (1) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility),
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (1) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (1) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or alloimmunisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various neuropathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L_{h4}.
 CC
 XX
 SQ Sequence 1139 BP; 290 A; 300 C; 283 G; 266 T; 0 other;
 Query Match 90.8%; Score 1067.4; DB 24; Length 1139;
 Best Local Similarity 97.5%; Pred. No. 1.3e-285;
 Matches 1103; Conservative 0; Mismatches 6; Indels 22; Gaps 1;
 QY 44 CCTCCAGGTTCTGGTGTGTAATGAATGAAGGCCCCAGATGCAAGATGCT 103
 DB 30 CCCACCGGTTCTGGTGTGTAATGAATGAAGGCCCCAGATGCAAGATGCT 89
 QY 104 GAAGGCTCCAGGCTGCTTCACTGCAAGCTGCCAGGGCTGGAAGCTCATCATG 163
 DB 90 GAAGGCTCCAGGCTGCTTCACTGCAAGCTGCCAGGGCTGGAAGCTCATCATG 149
 QY 164 GAGCTCAGTGAATGAGTGGTCTAAGCGTGAAGGCCATGAGGCCCATCACCATTGA 223
 DB 150 GAGCTCAGTGAATGAGTGGTCTAAGCGTGAAGGCCATGAGGCCCATCACCATTGA 209
 QY 224 CCGCTTCACTCTCAGAGTGAAGTGAAGGCCAGGAATTCCTCGAGATGATCATCA 283
 DB 210 CCGCTTCACTCTCAGAGTGAAGTGAAGGCCAGGAATTCCTCGAGATGATCATCA 269
 QY 284 CAATGTGAGGCCAGTGAATGGGGGAACATGATGCAAGCTCGAGAACTGCTGCA 343

PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 (AMGB-) AMGEN INC.
 PI Welcher AA, Sarmiento UM, Schultz HU, Chute HT;
 DR WPI: 2002-130881/17.
 XX P-PSDB; A075540.
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 XX for diagnosing, preventing and treating reproductive, immune and
 XX proliferative disorders, e.g. cancer and arteriosclerosis
 PS Claim 1; Fig 1; 135pp: English.
 XX
 CC The invention relates to an isolated B7-like (B7-L) polypeptide (I),
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of hematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or all sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L_{h1}.
 XX
 SO Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other;
 Query Match 100.0%; Score 1175; DB 24; Length 1175;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 181 TGTGCTAAGCGTCAGGCCCCATGAGCCCATCATCACCACATGACCGTCTCAGCA 240
 QY 241 GGTACGACCGAGGGGGGAACTTCACTCGAGATGATCATCACAATGTGAGCCAGTG 300
 DB 241 GGTACGACCGAGGGGGGAACTTCACTCGAGATGATCATCACAATGTGAGCCAGTG 300
 QY 301 ATTGCGGAGACATGAGATGACAGCTCCAGAACAGTGGCTGATGATCTGCTTACCTTA 360
 DB 301 ATTGCGGAGACATGAGATGACAGCTCCAGAACAGTGGCTGATGATCTGCTTACCTTA 360
 QY 361 CCGTCAAGTATGAGGAGAGCTGTTTCATTCCTCCAGTGTAACTGTTGTCGAGAGATG 420
 DB 361 CCGTCAAGTATGAGGAGAGCTGTTTCATTCCTCCAGTGTAACTGTTGTCGAGAGATG 420
 QY 421 AACCTTGAAGTACTGTTCTACCTCCACTGAGACCGGCTCCGAGATATTTCTGCG 480
 DB 421 AACCTTGAAGTACTGTTCTACCTCCACTGAGACCGGCTCCGAGATATTTCTGCG 480
 QY 481 AGCTGCTCTCTGCTGACAGCCATTCAGCTTATTTGTTCCGAGCCGACGACTTC 540
 DB 481 AGCTGCTCTCTGCTGACAGCCATTCAGCTTATTTGTTCCGAGCCGACGACTTC 540
 QY 541 AAGTGCAGTGCATTCCTGGGCTGACCCCAAGAGCAATGGAGACTTGTGCTG 600
 DB 541 AAGTGCAGTGCATTCCTGGGCTGACCCCAAGAGCAATGGAGACTTGTGCTG 600
 QY 601 CTACCTGGAAGACCTGGAAGGCCGCAAGTCTGCAACTGTAATCTCAGTGTGCT 660
 DB 601 CTACCTGGAAGACCTGGAAGGCCGCAAGTCTGCAACTGTAATCTCAGTGTGCT 660
 QY 661 GTCCCAAGACCTGAGAGTGTATTAATTTCCAGGTGTATTAATTAAGTGTACCGAGT 720
 DB 661 GTCCCAAGACCTGAGAGTGTATTAATTTCCAGGTGTATTAATTAAGTGTACCGAGT 720
 QY 721 TAGGTTTTCATTCGCTACTGTTGGGCAAGTTGAGCTGAGTACGACGACCAATGCTTC 780
 DB 721 TAGGTTTTCATTCGCTACTGTTGGGCAAGTTGAGCTGAGTACGACGACCAATGCTTC 780
 QY 781 TGAAGCGGAGCTGATCTTACATATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 781 TGAAGCGGAGCTGATCTTACATATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 841 GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 TTCAAAAGAAATCTGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 960
 DB 901 TTCAAAAGAAATCTGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 960
 QY 961 AAAACTCCGGCTACAAATTCATGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1020
 DB 961 AAAACTCCGGCTACAAATTCATGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1020
 QY 1021 AATCTGTGATTCAGTGTCTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080
 DB 1021 AATCTGTGATTCAGTGTCTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080
 QY 1081 GGGCTGATCAAGCTCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
 DB 1081 GGGCTGATCAAGCTCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
 QY 1141 GTCTGAGAAAGTCAAGTAATACACTGATGATAG 1175
 DB 1141 GTCTGAGAAAGTCAAGTAATACACTGATGATAG 1175
 RESULT 2
 ABK13030
 ID ABK13030 standard; cDNA; 1240 BP.
 XX
 AC ABK13030;
 XX


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9990
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5063 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA/DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 864..1349
FEATURE:
NAME/KEY: CDS
LOCATION: 1944..3370
FEATURE:
NAME/KEY: CDS
LOCATION: 3439..3736
US-08-185-432-1

```

```

Query Match          3.5%; Score 41; DB 1; Length 5063;
Best Local Similarity 66.3%; Pred. No. 0.11;
Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

```

QY 808 GCTGCTCTCTGCGCGCGCTGTTGGTGCACACTGCTGCCGTTGTTCT 867
      |||||
DB 2988 GCTGCTCTCTGCGCGCTGTTGGTGCACACTGCTGCCGTTGTTCT 867
      |||||
QY 868 GCTGTAGAGAGAGAGAGATTTCGTATT 896
      |||||
DB 2928 GCTGTGCTGTCGCGGATGCAGTGTATT 2900
      |||||

```

```

RESULT 15
US-08-479-913E-1/c
; Sequence 1, Application US/08479913E
; Patent No. 6416998
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Bert W.
; APPLICANT: Tsai, Ming-Jer
; APPLICANT: Ledebur, Harry C.
; APPLICANT: Kittle, Joseph D.
; TITLE OF INVENTION: MODIFIED STEROID HORMONES FOR GENE THERAPY AND METHODS FOR THEIR
; FILE REFERENCE: 212/133
; CURRENT APPLICATION NUMBER: US/08/479,913E
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 07/939,246
; PRIOR FILING DATE: 1992-09-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 1
; LENGTH: 6177
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Plasmid pGR0403R
US-08-479-913E-1

```

```

Query Match          3.5%; Score 40.8; DB 4; Length 6177;
Best Local Similarity 75.0%; Pred. No. 0.14;
Matches 51; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY 808 GCTGCTCTCTGCGCGCGCTGTTGGTGCACACTGCTGCCGTTGTTCT 867
      |||||
DB 4785 GCTGCTCTCTGCGCGCTGTTGGTGCACACTGCTGCCGTTGTTCT 867
      |||||
QY 868 GCTGTAGA 875
      |||||
DB 4725 GCTGCACA 4718
      |||||

```

```

Search completed: April 29, 2003, 10:49:16
Job time : 88.1906 secs

```

```

TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 3771 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-185-432-23

Query Match 3.5%; Score 41; DB 1; Length 3771;
Best Local Similarity 66.3%; Pred.No. 0.096;
Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0.

QY 808 GCTGCTGCTCTGCTCCCGCCGCTGTTGTTGGCTGCAACTGCTGCTGCCGTTGTTCTTC 867
      ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1897 GCTGCTGCTGCTGCGCTGTGCTGTGCTGCATCTGCAGCTGTTGCTGTTGTTGCTGT 1956

QY 868 GCTGTAGAGAAAAGAGATTCGATT 896
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1957 GCTGTGCTGTGCGGATGCATGCTGAAT 1985

      US-08-185-432-1/c
      ; Sequence 1, Application US/08185432
      ; Patent No. 5750652
      ; GENERAL INFORMATION:
      ; APPLICANT: Artavanis-Tsakonas, Spyridon
      ; APPLICANT: Busseau, Isabelle
      ; APPLICANT: Diederich, Robert J.
      ; APPLICANT: Xu, Tian
      ; APPLICANT: Matsuno, Kenji
      ; TITLE OF INVENTION: DEUTEX PROTEINS, NUCLEIC ACIDS, AND
      ; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
      ; NUMBER OF SEQUENCES: 23
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: PENNIE & EDMONDS
      ; STREET: 1155 Avenue of the Americas
      ; CITY: New York
      ; STATE: New York
      ; COUNTRY: U.S.A.
      ; ZIP: 10036-2711

      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30

```

Query Match	3.58; Score 41.6; DB 4; Length 203;
-------------	-------------------------------------

154 TGCTGCTGCTGATGTTGCTGCTGTTG 129

846 TECTGCTGCCCTTGTGTCTGCTG 871

154 TCGCTGCTGATGTTGCTGCTGTG 129

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1626
: US-08-328-809-7
:
: Query Match 3.6%; Score 42.6; DB 1; Length 3292;
: Best Local Similarity 64.9%; Pred. No. 0.031;
: Matches 63; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
:
: Oy 804 ATACGCTGCTGCTGCTGCCGCCGTCGTCGTGTGTGCTGCACACTGCTGCTGCCGTGTGTGT 863
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 818 AGATGATGATGATGTTGTTGTGAAGTGTGTGTTGTTGTGCAACTGATGTTGTGATGAGC 759
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Oy 864 TTCTGCTGACAGAAAAGAGATTGCTAFTTCAT 900
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 758 TGCTGCTGTTGCTGCACACTGCTGTTGTTGTTAATT 722
:
: RESULT 8
: PCT-US92-11107-12/c
: Sequence 12, Application PC/TUS9211107
: GENERAL INFORMATION:
: APPLICANT: Donahue, Brian A.
: APPLICANT: Toney, Jeffrey H.
: APPLICANT: Bruhn, Suzanne L.
: APPLICANT: Pil, Pieter M.
: APPLICANT: Brown, Steven
: APPLICANT: Kellelt, Patti
: APPLICANT: Essigmann, John M.
: APPLICANT: Léopard, Stephen J.
: TITLE OF INVENTION: DNA Structure Specific Recognition
: TITLE OF INVENTION: Protein and Uses Therefor
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: 2 Millitia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/11107
: FILING DATE: 19921218
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/539,906
: FILING DATE: 18-JUN-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,227
: REFERENCE/DOCKET NUMBER: MIT-4787AAA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 617-861-9540
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3292 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: Double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Saccharomyces cerevisiae
: IMMEDIATE SOURCE:
: CLONE: lambda yPt
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1626

```

ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/258,442
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/539,906
 FILING DATE: 18-JUN-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: MIT-4787AAA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ. ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3292 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

```

Query Match          4.1%; Score 48.2; DB 2; Length 325;
Best Local Similarity 57.8%; Pred. No. 0.00023;
Matches 100; Conservative 2; Mismatches 70; Indels 1; Gaps 1

QY 792 TGACCTTTCACAAATACGCTGCTGCTGCTGCCGCGTCGTTGTTGGCTGCACACTGCTGC 851
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 TGTCCTGATGAGTCCSCCTGCTGCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGC 151

QY 852 TGCCTGTGTTGTTTCTGCTGTAGAA-GAAAAGAGACATTGCTATTCATTTCAAAAGAA 910
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 TGCCTGTGCTCTTTTCTGCTGCTGCTGGAACACTTCAAAAGTGAAGTATTTATTTAAAAACA 91

QY 911 ATCTGAAAAGAGACAAACAAAGAAAGTGAAGACAGAAGATGGAATGAAA 963
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 90 AAACCTTAAAGATAATATACACCATGAGAAAATATTCATTAAGGAAAATACA 38

RESULT 3
US-08-531-927B-1/c
; Sequence 1, Application US/08531927B
; Patent No. 5840491
; GENERAL INFORMATION:
; APPLICANT: Kakizuka, Akira
; TITLE OF INVENTION: DNA Sequence Encoding the Machado-Joseph
; Patent No. 5840491
; TITLE OF INVENTION: Disease Gene and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
;

```

	Query Match	3.6%	Score 42.8	DB 2	Length 1776
	Best Local Similarity	61.8%	Pred. No. 0.02		
	Matches 68	Conservative	0	Mismatches 42	Indels 0
				Gaps 0	
QY	808	GCCTGACGCTGCTCCGCCCTGCTTTTGGCTGCAATGCTGCTGCGCTGTTTCTTCT	867		
DB	975	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTT	916		
QY	868	GCCTGTAAGAAAAGAGATTTCTGATTCATTTTCAAAAGAAATTCGAA	917		
DB	915	GCCTGCTGTTTTTTCAAAGTAGGCTTCGCTCTCTTCCGAAGCTCTTTCGAA	866		

RESULT 4
 US-09-041-886-12/c
 ; Sequence 12, Application US/09041886
 ; Patent No. 6235872
 ; GENERAL INFORMATION:
 ; APPLICANT: Bredesen, Dale E.
 ; APPLICANT: Rabizadeh, Sharoz
 ; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
 ; TITLE OF INVENTION: Polypeptides and Methods of Use
 ; NUMBER OF SEQUENCES: 72
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell & Flores LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 92122
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/041,886
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-LJ 2626
 ; TELECOMMUNICATION INFORMATION:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.


```

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weisheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeonica-X-1
FILE REFERENCE: GENE EXPRESSION ANALYSIS BY MICROARRAY
CURRENT FILING DATE: 2001-05-23
PCT APPLICATION NUMBER: US/09/864,761
PCT FILING DATE: 2000-02-04
PCT APPLICATION NUMBER: US 60/207,456
PCT FILING DATE: 2000-05-26
PCT APPLICATION NUMBER: US 09/632,366
PCT FILING DATE: 2000-08-03
PCT APPLICATION NUMBER: GB 24263.6
PCT FILING DATE: 2000-10-04
PCT APPLICATION NUMBER: US 60/236,359
PCT FILING DATE: 2000-09-27
PCT APPLICATION NUMBER: PCT/US01/00666
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00667
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00664
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00669
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00665
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00668
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00663
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00662
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00661
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: PCT/US01/00670
PCT FILING DATE: 2001-01-30
PCT APPLICATION NUMBER: US 60/234,687
PCT FILING DATE: 2000-09-21
PCT APPLICATION NUMBER: US 09/608,408
PCT FILING DATE: 2000-06-30
PCT APPLICATION NUMBER: US 09/774,203
PCT FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 18923
LENGTH: 293
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009954.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 47
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 53
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 69
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 21
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 45
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 21
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 33
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 21
OTHER INFORMATION: NT HIT: Z22828.1, EVALUE 9.00e-59
OTHER INFORMATION: SWISSPROT HIT: P53360, EVALUE 4.00e-15
OTHER INFORMATION: EST_HUMAN HIT: AW409956.1, EVALUE 1.00e-58
US-09-864-761-18923
Query Match 3.7% Score 43.2; DB 10; Length 293;
Best Local Similarity 79.7%; Pred. No. 0.005;
Matches 51; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

[illegible]

PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988

Patent No. US20020048763A1

RESULT 13
 JS-09-664-761-18923/C
 Sequence 18923, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:

XX

XX

XX

XX 28-JUN-2001; 2001WO-US20719.
 XX
 PF 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 DR WPI; 2002-130881/17.
 XX P-PSDB; AAU75545.
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 1; Fig 6; 135pp; English.

XX The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility)
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or alloimmunisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions, and for
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of mouse B7-L_{m2}.
 XX

SQ Sequence 895 BP; 250 A; 223 C; 212 G; 210 T; 0 other;

Query Match 15.0%; Score 185.6; DB 24; Length 895;
 Best Local Similarity 73.8%; Pred. No. 5.8e-41;
 Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 115 AGTCTGGGTCTGTAATGAGTATAGAGGCCCCAGATGCAAGAGTCTGGAAGG 174
 DB 109 AGCTTCGGGATCCAGTATACATAGAGGCTCTCAGAAATGTAACAGTCTTAAGA 168
 DB 175 CTCCAGGGCTCGTCAACGACCGCTCTCCAGGCTGGAAGCTCATATGAGGCTCT 234
 DB 169 CTGAGAGGCTCACTTAAGCTGACGCTGACTCAAGGCTTGAAGCTTCTATGTGACTCT 228
 QY 235 CAGTGAACATGCTGCTAAGCGTCAGGCCATGAGCCATCATACCAAGACCGGCT 294
 DB 229 TAACCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288

QY 295 CACCTTCAGAGGTACAGACAGGGCGGAGCACTTACCTCGAGATGATATCCCAATGT 354
 DB 289 CACCTATGACAGTTACACAGCACTGACAGCTTCACTCGAGATGATATCATGATGT 348
 QY 355 GAGCCAGGATGATTCGGGGGAACATGACATGACAGCTTCCAGAACAGTGGCTGATGATC 414
 DB 349 GCAGCCAGGATGATTCGGGGATTCGTCAGATGACAGCTTCCAGAACAGGATGATGATC 408
 QY 415 TGCTTACCTTACCGTCCAG 434
 DB 409 TGCTTACCTTACCGTCCAG 428

RESULT 10
 ABK10239/C
 ID ABK10239 standard; DNA; 277 BP.
 XX
 AC ABK10239;
 XX
 DT 20-MAY-2002 (first entry)
 XX
 DE Trinucleotide repeat sequence #2.

XX Trinucleotide repeat; fragile X syndrome; ds; DRPLA;
 KW spinocerebellar ataxia type III; Marfan syndrome;
 KW hereditary hypertrophic cardiomyopathy; neuropsychiatric;
 KW dentatorubral and pallidolysian atrophy;
 KW spinocerebellar ataxia; X-linked spinobulbar atrophy.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT repeat_region 152..194
 FT /tag= a
 FT /rpt_type= TANDEM
 FT 152..154
 FT /tag= b
 FT /note= "CAG type repeat"
 FT 198..222
 FT /tag= a
 FT /rpt_type= TANDEM
 FT 198..200
 FT /tag= b
 FT /note= "CAG type repeat"

XX Homo sapiens.
 OS KR2000003004-A.
 XX
 PN 15-JAN-2000.
 PD
 XX 25-JUN-1998; 98KR-0024064.
 PF 25-JUN-1998; 98KR-0024064.
 PR 25-JUN-1998; 98KR-0024064.
 XX
 PA (SMSU) SAMSUNG MEDICAL CENT.
 PA (JIND/) JIN D. G.

XX JIN DG;
 DR WPI; 2000-662424/64.

XX Plasmid vector arraying neuropsychiatric gene more than copy number
 PT containing trinucleotide repeats its use thereof -
 XX
 PS Disclosure; Page 12; 23pp; Korean.

XX This invention relates to the use of a plasmid vector array
 CC containing a neuropsychiatric gene containing trinucleotide repeats.
 CC Trinucleotide repeats and their expansion are known to be involved
 CC in a number of diseases including fragile X syndrome, spinocerebellar
 CC ataxia type III, hereditary hypertrophic, cardiomyopathy, dentatorubral

ID	ABK13034	standard; cDNA; 754 BP.
XX	ABK13034;	
XX	23-APR-2002	(first entry)
XX	DNA encoding mouse B7-1-like protein, B7-L _{m3} .	
XX	Mouse; B7-1-like protein; B7-1; antiinfectility; gynaecological;	
KW	antitumour; cyclostatic; immunosuppressive; antiarthritis; antirheumatic;	
KW	antihypertensive; dermatologic; antipsoriatic; neuroprotective;	
KW	antidiabetic; haemostatic; antihypoid; antulcer; antiallergic;	
KW	antiaesthetic; nephrotoxic; antibacterial; vitruclide; tumour; cancer;	
KW	reproductive disorder; graft versus host disease; autoimmune disease;	
KW	toxic shock syndrome; allergy; nephropathy; skin disorder;	
KW	endocrinopathy; lymphoproliferative disorder; gene; ss.	
XX	mus musculus.	
OS	Key	Location/Qualifiers
PH	CDS	53..724
FT		/*tag= a
FT		/product= "B7-1-like protein, B7-L _{m3} "
XX	W0200200710-A2.	
PM	03-JAN-2002.	
PD	28-JUN-2001; 2001WO-US20719.	
XX	28-JUN-2000; 2000US-214512P.	
XX	28-NOV-2000; 2000US-0729264.	
PR	(AMGE-) AMGEN INC.	
XX	Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;	
PI	WPI: 2002-130881/17.	
DR	P-PSDB: AMU75546.	
XX	New B7-1-like polypeptides, polynucleotides and their modulators, useful	
PT	for diagnosing, preventing and treating reproductive, immune and	
PT	proliferative disorders, e.g. cancer and arteriosclerosis	
XX	Claim 1; Fig 7; 135pp; English.	
PS	The invention relates to an isolated B7-1-like (B7-L) polypeptide (I);	
XX	The polypeptide, polynucleotide encoding it and antibody against (I) are	
CC	useful for treating B7-1-like polypeptide-related disease, disorders or	
CC	conditions including reproductive disorders (e.g. infertility,	
CC	miscarriage, preterm labour and delivery and endometriosis) and	
CC	proliferative disorders. Antibodies, soluble proteins comprising	
CC	extracellular domains and other regulators of B7-L polypeptides are	
CC	useful for enhancing the immune response to tumours. (I) plays a role in	
CC	growth and maintenance of cancer cells based on the observation of	
CC	seminal vesicle hyperplasia in transgenic mice overexpressing B7-L	
CC	polypeptide. Hence modulators of (I) are useful for the treatment of	
CC	cancer including seminal vesicle cancer, lung, brain, breast, ovarian,	
CC	testicular cancer and cancers of haematopoietic system. B7-L polypeptide	
CC	pathway can be manipulated to regulate cytotoxic T-lymphocyte response	
CC	in allograft transplantation, graft versus host disease, T-cell	
CC	dependent B-cell mediated diseases and autoimmune diseases. B7-L	
CC	molecules are useful for alleviating the symptoms associated with	
CC	diseases involving chronic immune cell dysfunction or to treat	
CC	autoimmune diseases such as systemic lupus erythematosus, Rheumatoid	
CC	arthritis, multiple sclerosis, diabetes, immune thrombocytopenic	
CC	purpura and psoriasis, chronic inflammatory disease such as	
CC	inflammatory bowel disease, (Crohn's disease and ulcerative colitis),	
CC	grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They	
CC	are also useful as immunosuppressive agents for bone marrow and organ	

Query Match	15.0%	Score 185.6	DB 24	Length 754
Best Local Similarity	73.8%	Pred. No. 5.3e-41		
Matches 236	Conservative	0	Mismatches 84	Indels 0
			Gaps	0
YY	115 AGGTTCTGGGTCGTGTAATGAAGTACATAGAGAGCCGCCAGAAATGCAACAGTCTCTGAGGG	174		
DB	109 AACTTCCGGATCCAGATTATACATCATAGAGGTCTCTCAGAAATGTAACAGTCTCTAAGGA	168		
YY	175 CTCGCCAGGCTCGCTTCAACTGACACGCTCCCAAGGCGTGAAGTCATATGNGGCTCT	234		
DB	169 CTCAGAGGCTCACTTCAACTGACACGCTGACTACGCGCTGAACTTCTATGTGGACTCT	228		
YY	235 CAGTGACATAGTGTGCTGAAGGCTCAGGCGCATGAGACCATCATACCAATGACCGCTT	294		
DB	229 TAACCAAAATGTGTGTCTGATCTGATCCACCAACGAGACCATCATACCAACACACCGCTT	288		
YY	295 CACCTTCAGAGGTACGACACGAGGCGGGAACCTTCACTCGGAGTATCATCCACATGT	354		
DB	289 CACCTTATGCAAGTTAAACAGACACTACAGCTTCACTCGGAGTATCATCCATGTGT	348		
YY	355 GGAGCCAGATGATTCGGGGAAATCATGAGACGCTCCAGAACAGTGGCCTGCATGATTC	414		
DB	349 GCAGCCCAAGTATCGGGATCCGTGCAATGACACCTGACAGAACAGCATGAGTGGATC	408		
YY	415 TGCTTACCTTACCGTCAAG	434		
DB	409 TGCTTCTCTCTCAGTCAAG	428		

RESULT 9

ID ABK13033

XX ABK13033 standard; cDNA; 895 BP.

XX

AC ABK13033;

XX

DT 23-APR-2002 (first entry)

XX

DE DNA encoding mouse B7-like protein, B7-L_m2.

XX

XX Mouse: B7-like protein; B7-L; antiinfectivity; gynaecological;

KW antitumor; cytosolic; immunosuppressive; antiatheritic; antirheumatic;

KW antiinflammatory; dermatological; antipsoriatic; neuroprotective;

KW antidiabetic; haemostatic; antihypoid; antiulcer; neuroallergic;

KW antistimatic; nephrotoxic; antibacterial; virucide; tumour; cancer;

KW reproductive disorder; graft versus host disease; autoimmune disease;

KW toxic shock syndrome; allergy; nephropathy; skin disorder;

KW endocrinopathy; lymphoproliferative disorder; gene; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 53..865

FT /*tag= a

FT /product= "B7-like protein, B7-L_m2"

XX

PN W0200200710-A2.

XX

PD 03-JAN-2002

QY 115 AGCTTGGGCTGCTGATATGAGTCAATGAAAGGCCCCAGATGCAACAGCTCTGAAGG 174
 DB 109 AGCTTCCGGATCCAGTATATGATAGAGAGGCTCTGCAATATGATACAGTCTTAAAGA 168
 QY 175 CTCCAGGCTGCTTCAACGTCACCTCTCCAGGGCTGGAAGCTCATCTATGGGCTCT 234
 DB 169 CTGAGAGGCTACTTCAAGTCACTGACCTGAGCTGAGGCTGAGCTTCTGAGACTCT 228
 QY 235 CAGTACATGGTGGTCTAAGCCGTCAGGAGCCATGAGAGCCATCATGACAGCGCTT 294
 DB 229 TAACCAATATGGTGGTCTGAGTCTACACACCAAGAGCCATCATGACCAACACCGCTT 288
 QY 295 CACCTTCAGAGGTACAGCAGGAGGCGAACTTACCTGGAGATGATCATCCCAATGT 354
 DB 289 CACCTATGCCAGTTACAAACAGCACTGACCTCATCTCGAGTGTGATCATCATGATGT 348
 QY 355 GAGACCGAGTATGGGGGGAACATGACATGACACCTCCAGAACAGTCCCTGATGATC 414
 DB 349 GACGCCAGTACTGCGGATCCGTCATGACAGCTGACAGACAGCCATGGGTTTGGATC 408
 QY 415 TGCTTACCTTACCTGTCACATGATGAGAGAGCTGTCATCCAGTGTATCTTGTAGT 474
 DB 409 TGCTTCTCTCATGATGATGATGAGAGCTGACCATCTTCTAGCAACACCTTATAGT 468
 QY 475 CGCTGAGATGAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 534
 DB 469 CACGAGGAGTGAACCTGTATGATGATGATGATGATGATGATGATGATGATGATG 528
 QY 535 TATTTCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
 DB 529 TATTTCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
 QY 595 CAGGACCTTCAAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGTGC 654
 DB 589 GGGCACTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
 QY 655 GACTTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
 DB 649 GACTTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
 QY 715 TGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
 DB 709 TGTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
 QY 775 TTTACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 834
 DB 745 TTTACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
 QY 835 CACCATGCTTCTGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
 DB 793 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
 QY 895 TTTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
 DB 844 TTTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
 QY 955 TCGTATTCATTTCAAAAGAAATCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 1014
 DB 900 CTGCAAAAGTGAAGCA-----ACAAAGCAGATCCGAGACAAAGTTAA 945
 QY 1015 AAGTGAATGAAATCTCGGGCTACATTCAGATGAAACAAAGAACACAGAAACCGCTTC 1074
 DB 946 AAGTGAATGAAATCTCGGGCTACATTCAGATGAAACAAAGAACACAGAAACCGCTTC 1005
 QY 1075 TCTGCCCTCCAAATCTGTAATTCAGATGCTGTAATTCAGATGCTGTAATTCAGATG 1134
 DB 1006 TCTGCCCTCCAAATCTGTAATTCAGATGCTGTAATTCAGATGCTGTAATTCAGATG 1059
 QY 1135 TCTGCCCTCCAAATCTGTAATTCAGATGCTGTAATTCAGATGCTGTAATTCAGATG 1194
 DB 1060 TCTGCCCTCCAAATCTGTAATTCAGATGCTGTAATTCAGATGCTGTAATTCAGATG 1119

QY 1195 TATCTGGCCAGTCTGAGAGAGTCAATATGATGATGATGATGATGATGATGATG 1239
 DB 1120 TATCATGCGCACTCTCTGAGAGAGTCAATATGATGATGATGATGATGATGATG 1164
 RESULT 7
 ID AAI36582
 AA136582 standard; DNA; 401 BP.
 AC AA136582;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #5268 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 PS Claim 25; SEQ ID No 5268; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 401 BP; 105 A; 98 C; 111 G; 87 T; 0 other;
 Query Match 15.2%; Score 188.8; DB 22; Length 401;
 Best Local Similarity 99.0%; Pred. No. 5e-42;
 Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 115 AGTTCCTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 174
 DB 210 AGTTCCTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 269
 QY 175 CTCCAGGCTGCTTCAACGTCACCTCTCCAGGGCTGGAAGCTCATCTATGGGCTCT 234
 DB 270 CTCCAGGCTGCTTCAACGTCACCTCTCCAGGGCTGGAAGCTCATCTATGGGCTCT 329
 QY 225 CAGTACATGGTGGTCTAAGCCGTCAGGAGCCATGAGAGCCATCATGACCAATGACCGCTT 294
 DB 330 CAGTACATGGTGGTCTAAGCCGTCAGGAGCCATGAGAGCCATCATGACCAATGACCGCTT 389
 QY 295 CACCTCTCAGAG 306
 DB 390 CACCTCTCAGAG 401

```

Db 241 AATCTTGATGCTGAGATGACCTTGTGAAGTTACTTGTCTACCTCCACACAGCGACC 300
      |||||||
Qy 524 CGGCTCCGCGATATTTCTGGAGCTCGCTCTCGGTAGCCATTCAAGTATATTTT 583
      |||||||
Db 301 CGGCTCCGCGATATTTCTGGAGCTCGCTCTCGGTAGCCATTCAAGTATATTTT 360
      |||||||
Qy 584 GTTCGCGAGCCCGACCTTCAAAATGAGAGACATCTGCTGACCCACACAGCC 643
      |||||||
Db 361 GTTCGCGAGCCCGACCTTCAAAATGAGAGACATCTGCTGACCCACACAGCC 420
      |||||||
Qy 644 AATGGAAGCTTGAAGCTGAGCTGAGAGAGCCCTAAGCCGCAAGTCTGCACT 703
      |||||||
Db 421 AATGGAAGCTTGAAGCTGAGCTGAGAGAGCCCTAAGCCGCAAGTCTGCACT 480
      |||||||
Qy 704 GTAATCTCACTGATGATGCTGCTCCCAAGACATGAGAGGTATTAATATCCAGGT 763
      |||||||
Db 481 GTAATCTCACTGATGATGCTGCTCCCAAGACATGAGAGGTATTAATATCCAGGT 540
      |||||||
Qy 764 GTATTTATCAAGTTACCGAGTTAGTTTTCATTTGCTGCTGAGGCGCAAGTGGACTT 823
      |||||||
Db 541 GTATTTATCAAGTTACCGAGTTAGTTTTCATTTGCTGCTGAGGCGCAAGTGGACTT 600
      |||||||
Qy 824 GGAATGACAGCAGCAGATGCTTTCAGAGCGAGCTGACTTTCATTAATAGCTGCTGCTGC 883
      |||||||
Db 601 GGAATGACAGCAGCAGATGCTTTCAGAGCGAGCTGACTTTCATTAATAGCTGCTGCTGC 660
      |||||||
Qy 884 TGCCGCGCTGTTGTTGTGGCTGCACTGCTGCTGCTGTTGTTGCTGCTGTAAGAA 943
      |||||||
Db 661 TGCCGCGCTGTTGTTGTGGCTGCACTGCTGCTGCTGTTGTTGCTGCTGTAAGAA 720
      |||||||
Qy 944 AAAAGAGATTTGCTGATTC----- 962
      |||||||
Db 721 AAAAGAGATTTGCTGATTC----- 962
      |||||||
Qy 963 ----- 962
      |||||||
Db 781 AATGGCGTCAATTTGCTGATTCATTTATGCTGCTGCTGAGCCCTCACTGTGAGA 840
      |||||||
Qy 963 -----AATTTCAAAAATAATCTGAAAAGAGAAACAAAGAAAGTGAAGCAG 1013
      |||||||
Db 841 CTGATTCATACCGTGGCAGCAAAATCTGAAAAGAGAAACAAAGAAAGTGAAGCAG 900
      |||||||
Qy 1014 AATGGAATGAAAAGTCCGCGCTACAAATCAATGAAAGAAAGACACAGAAACCGCTT 1073
      |||||||
Db 901 AATGGAATGAAAAGTCCGCGCTACAAATCAATGAAAGAAAGACACAGAAACCGCTT 960
      |||||||
Qy 1074 CTCTCCCTCCCAATCTCTGATTCAGTATCTGAAACAAAGAAAGACAGTGTGGCC 1133
      |||||||
Db 961 CTCTCCCTCCCAATCTCTGATTCAGTATCTGAAACAAAGAAAGACAGTGTGGCC 1020
      |||||||
Qy 1134 CTCTCCAGCAGCGGCTGATCAAGT 1159
      |||||||
Db 1021 CTCTCCAGCAGCAAGACATAGACCT 1046
      |||||||

```

RESULT 6
ABK13032
ID ABK13032 standard; cDNA; 1195 BP.

XX ABK13032;
XX 23-APR-2002 (first entry)

DE DNA encoding mouse B7-1-like protein, B7-L_{m1}.

XX Mouse: B7-1-like protein; B7-L_{m1} antineoplastic; gynecological;
XX antitumor; cytostatic; immunosuppressive; antiarthritic; antirheumatic;
XX antiinflammatory; dermatological; antipsoriatic; neuroprotective;
XX antidiabetic; haemostatic; antihypertensive; antitumor; antiallergic;
XX antineoplastic; nephrotoxic; antibacterial; virucide; tumour; cancer;
XX reproductive disorder; graft versus host disease; autoimmune disease;
XX toxic shock syndrome; allergy; neuropathy; skin disorder;

KW endocrinopathy; lymphoproliferative disorder; gene; ss.
XX Mus musculus.
OS
XX Key Location/Qualifiers
FH 53..1165
FT /"tag" a
FT /product= "B7-1-like protein, B7-L_{m1}"
XX
XX W020020710-A2.
XX
XX 03-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US20719.
XX
XX 28-JUN-2000; 2000US-214512P.
XX 28-NOV-2000; 2000US-0729264.
XX
XX (AMGE-) AMGEN INC.
XX
XX Weicher AA, Sarmiento UM, Schultz HJ, Chute HT;
XX WPL: 2002-130881/17.
XX P-PSDB: AA075544.
XX
XX New B7-1-like polypeptides, polynucleotides and their modulators, useful
XX for diagnosing, preventing and treating reproductive, immune and
XX proliferative disorders, e.g. cancer and arteriosclerosis
XX
XX Claim 1; Fig 5; 135pp; English.

XX The invention relates to an isolated B7-1-like (B7-L) polypeptide (I).
XX The polypeptide, B7-1-like polypeptide-related disease, disorders or
XX useful for treating B7-1-like polypeptide-related disease, disorders or
XX conditions including reproductive disorders (e.g. infertility,
XX miscarriage, preterm labour and delivery and endometriosis) and
XX proliferative disorders. Antibodies, soluble proteins comprising
XX extracellular domains and other regulators of B7-L polypeptides are
XX useful for enhancing the immune response to tumours. (I) plays a role in
XX growth and maintenance of cancer cells based on the observation of
XX seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
XX polypeptide. Hence modulators of (I) are useful for the treatment of
XX cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
XX testicular cancer and cancers of haematopoietic system. B7-L polypeptide
XX pathway can be manipulated to regulate cytotoxic T-lymphocyte response
XX in allograft transplantation, graft versus host disease, T-cell
XX dependent B-cell mediated diseases and autoimmune diseases. B7-L
XX molecules are useful for alleviating the symptoms associated with
XX diseases involving chronic immune cell dysfunction or to treat
XX autoimmune diseases such as systemic lupus erythematosus, rheumatoid
XX arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
XX purpura and psoriasis, chronic inflammatory disease such as
XX inflammatory bowel disease (Crohn's disease and ulcerative colitis),
XX Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
XX are also useful as immunosuppressive agents for bone marrow and organ
XX transplantation or to prolong graft survival. B7-L molecules are also
XX useful for diagnosis and treatment of diseases involving abnormal cell
XX proliferation, including arteriosclerosis and vascular stenosis.
XX Antagonists of B7-L polypeptides are useful for alleviation of toxic
XX shock syndrome or allosensitisation due to blood transfusions, and for
XX treatment of allergy, asthma and hypersensitivity reactions,
XX nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
XX pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
XX (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
XX thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
XX lymphoproliferative disorders such as multiple myeloma. The present
XX sequence represents the coding sequence of mouse B7-L_{m1}.

XX Sequence 1195 BP; 313 A; 304 C; 291 G; 287 T; 0 other;

XX Query Match 30.7%; Score 380.6; DB 24; Length 1195;
XX Best Local Similarity 62.4%; Pred. No. 7e-95;
XX Matches 702; Conservative 0; Mismatches 354; Indels 69; Gaps 4;

The invention relates to an isolated B7-1like (B7-1), polypeptide (I). The polypeptide, polynucleotide encoding it and antibody against (I) are useful for treating B7-1like polypeptide-related disease, disorders or conditions including reproductive disorders (e.g., infertility, miscarriage, preterm labour and delivery and endometriosis) and

042

ID ABK13028 standard; cDNA: 1175 BP.
 XX
 AC ABK13028;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA encoding human B7-1-like protein, B7-L.h1.
 XX
 KM Human; B7-1-like protein; B7-L; antifertility; gynaecological;
 KM antitumour; cytostatic; immunosuppressive; antiarthritic; antineumatic;
 KM antiinflammatory; dermatological; antipsoriatic; neuroprotective;
 KM antidiabetic; haemostatic; antihypertoid; antileucic; antiallergic;
 KM antiasclerotic; nephroprotective; antibacterial; virucide; tumour; cancer;
 KM reproductive disorder; graft versus host disease; autoimmune disease;
 KM toxic shock syndrome; allergy; neuropathy; skin disorder;
 KM endocrinopathy; lymphoproliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 27..1175
 FT /tag= a
 FT /product= "B7-1-like protein, B7-L.h1"
 XX
 PD MO200200710-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US20719.
 XX
 PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 DR MPI: 2002-130881/17.
 XX P-PSDB: AAT75540.
 DR
 XX
 PT New B7-1-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 PS Claim 1; Fig 1; 135pp: English.
 XX
 CC The invention relates to an isolated B7-1-like (B7-L) polypeptide (1),
 CC the polypeptide, polynucleotide encoding it and antibody against (1) are
 CC useful for treating B7-1-like polypeptide-related disease, disorders or
 CC conditions involving reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (1) plays a role in
 CC growth and maintenance of cancer cells based on the observation of
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (1) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC dependent B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or to prolong graft survival. B7-L molecules are also
 CC useful for diagnosis and treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.

CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allosensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions.
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various neuropathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L.h1.
 XX
 SQ Sequence 1175 BP; 295 A; 307 C; 286 G; 287 T; 0 other:
 XX
 Query Match 90.6%; Score 1123.4; DB 24; Length 1175;
 Best Local Similarity 99.9%; Pred. No. 1.6e-300;
 Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 116 GGTTCGGTCTGCTGAATGAATGAAAGGCCCCAGATGCACAGCTCTGAAGGCG 175
 DB 51 GGTTCGGTCTGCTGAATGAATGAAAGGCCCCAGATGCACAGCTCTGAAGGCG 110
 QY 176 TCCGAGGCTGCTCAACTGACCGCTCCGAGGGCTGAAGCTCATCTGCGCTTC 235
 DB 111 TCCGAGGCTGCTCAACTGACCGCTCCGAGGGCTGAAGCTCATCTGCGCTTC 170
 QY 236 AGTACATGCTGCTGCTAAGGCTGAGGCCCATGAGCCATCATCACCAGTACCGCTTC 235
 DB 171 AGTACATGCTGCTGCTAAGGCTGAGGCCCATGAGCCATCATCACCAGTACCGCTTC 230
 QY 296 ACCCTCAGAGGTACGACACGAGGGGAGACTTCACCTCGAGATGATCCACAAATGTG 355
 DB 221 ACCCTCAGAGGTACGACACGAGGGGAGACTTCACCTCGAGATGATCCACAAATGTG 290
 QY 356 GAGCCAGTGAATGCGGGGAGACATGATGACGCTCCAGAACTGCTGCATGATCT 415
 DB 291 GAGCCAGTGAATGCGGGGAGACATGATGACGCTCCAGAACTGCTGCATGATCT 350
 QY 416 GCTTACCTTACCGTCAAGATTATGAGAGAGCTGTTCATTCCTGATTTATCTTGTAGTC 475
 DB 351 GCTTACCTTACCGTCAAGATTATGAGAGAGCTGTTCATTCCTGATTTATCTTGTAGTC 410
 QY 476 GCTGGAATGAACCTTGGAAGTACTGTCTACCTCAGACCTGAGCCGCTCCCGAT 535
 DB 411 GCTGGAATGAACCTTGGAAGTACTGTCTACCTCAGACCTGAGCCGCTCCCGAT 470
 QY 536 ATTTCCTGGAGCTCGCTCTCTGCTGACGACCATTCAGATTAATTTGTTCCGAGCC 595
 DB 471 ATTTCCTGGAGCTCGCTCTCTGCTGACGACCATTCAGATTAATTTGTTCCGAGCC 530
 QY 596 AGCGACCTTCAAGTGCAGTGCATCTGCTGACCCGACCCAGCAATGGGACTTTG 655
 DB 531 AGCGACCTTCAAGTGCAGTGCATCTGCTGACCCGACCCAGCAATGGGACTTTG 590
 QY 656 ACTTCGCGGCTACTGGAAGAGCTGGAAGGCCGACGCTGCAACTGTAATCTCACT 715
 DB 591 ACTTCGCGGCTACTGGAAGAGCTGGAAGGCCGACGCTGCAACTGTAATCTCACT 650
 QY 716 GTGATTCGGTCTCCCAAGACAGTGAAGTGTATTAATTCAGTGTATCAACT 775
 DB 651 GTGATTCGGTCTCCCAAGACAGTGAAGTGTATTAATTCAGTGTATCAACT 710
 QY 776 TTACGAGTGAATGTTTTCATGCTTCTGAGGCAAGTTGGACTGGACTGACAGC 835
 DB 711 TTACGAGTGAATGTTTTCATGCTTCTGAGGCAAGTTGGACTGGACTGACAGC 770
 QY 836 ACCATGCTTTCAGACGCGACGATGATCTTACATACGCTGCTGCTGCGCGCTGT 895
 DB 771 ACCATGCTTTCAGACGCGACGATGATCTTACATACGCTGCTGCTGCGCGCTGT 830
 QY 896 TGTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 DB 831 TGTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
 QY 956 CGTATTCATTTCAAAAGAAATCTGAAAAGAGAAAGCAAAAGAAAGTGAACAGAA 1015

PR 28-JUN-2000; 2000US-214512P.
 PR 28-NOV-2000; 2000US-0729264.
 XX (AMGE-) AMGEN INC.
 XX
 XX
 PI Welcher AA, Sarmiento UM, Schultz HJ, Chute HT;
 DR WPI: 2002-130881/17.
 DR P-PSDB: MAU75542.
 XX
 XX
 PT New B7-like polypeptides, polynucleotides and their modulators, useful
 PT for diagnosing, preventing and treating reproductive, immune and
 PT proliferative disorders, e.g. cancer and arteriosclerosis
 XX
 XX
 PS Claim 1: Fig 3; 135pp; English.

XX The invention relates to an isolated B7-like (B7-L) polypeptide (I).
 CC The polypeptide, polynucleotide encoding it and antibody against (I) are
 CC useful for treating B7-like polypeptide-related disease, disorders or
 CC conditions including reproductive disorders (e.g. infertility,
 CC miscarriage, preterm labour and delivery and endometriosis) and
 CC proliferative disorders. Antibodies, soluble proteins comprising
 CC extracellular domains and other regulators of B7-L polypeptides are
 CC useful for enhancing the immune response to tumours. (I) plays a role in
 CC growth and maintenance of the immune response to tumours. (I) plays a role in
 CC seminal vesicle hyperplasia in transgenic mice overexpressing B7-L
 CC polypeptide. Hence modulators of (I) are useful for the treatment of
 CC cancer including seminal vesicle cancer, lung, brain, breast, ovarian,
 CC testicular cancer and cancers of haematopoietic system. B7-L polypeptide
 CC pathway can be manipulated to regulate cytotoxic T-lymphocyte response
 CC in allograft transplantation, graft versus host disease, T-cell
 CC mediated B-cell mediated diseases and autoimmune diseases. B7-L
 CC molecules are useful for alleviating the symptoms associated with
 CC diseases involving chronic immune cell dysfunction or to treat
 CC autoimmune diseases such as systemic lupus erythematosus, rheumatoid
 CC arthritis, multiple sclerosis, diabetes, immune thrombocytopenic
 CC purpura and psoriasis, chronic inflammatory disease such as
 CC inflammatory bowel disease (Crohn's disease and ulcerative colitis),
 CC Grave's disease, Hashimoto's thyroiditis and diabetes mellitus. They
 CC are also useful as immunosuppressive agents for bone marrow and organ
 CC transplantation or treatment of diseases involving abnormal cell
 CC proliferation, including arteriosclerosis and vascular restenosis.
 CC Antagonists of B7-L polypeptides are useful for alleviation of toxic
 CC shock syndrome or allo sensitisation due to blood transfusions, and for
 CC treatment of allergy, asthma and hypersensitivity reactions,
 CC nephropathies (e.g. glomerulonephritis), skin disorders (pemphigus and
 CC pemphigoid), endocrinopathies (Grave's disease), various pneumopathies
 CC (extrinsic alveolitis), vasculopathies, coeliac disease, anaemia,
 CC thrombocytopenias, Guillain-Barre syndrome and myasthenia gravis, and
 CC lymphoproliferative disorders such as multiple myeloma. The present
 CC sequence represents the coding sequence of human B7-L_{H3}.
 XX

Sequence 1240 BP; 318 A; 319 C; 305 G; 298 T; 0 other;

Query Match 100.0%; Score 1240; DB 24; Length 1240;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTGATGTCAGCCAGCAAGTGTGATGATCTTCTAGGCTCCATACAAACACCAT 60
 DB 1 AGGTGATGTCAGCCAGCAAGTGTGATGATCTTCTAGGCTCCATACAAACACCAT 60
 QY 61 AACCTGTCCTTGAAGCAATGGAAGGCAATTTGTCAGGCTCCAGAGCTTACGTTCC 120
 DB 61 AACCTGTCCTTGAAGCAATGGAAGGCAATTTGTCAGGCTCCAGAGCTTACGTTCC 120
 QY 121 TGGGTCTGTAATGAACTATAGAGGCCCCAGAAATGCAACAGTCTGAAAGGCTCCCA 180
 DB 121 TGGGTCTGTAATGAACTATAGAGGCCCCAGAAATGCAACAGTCTGAAAGGCTCCCA 180
 QY 181 GGCCTGCTTCAAGTCAAGCGCTCCAGAGGCTGGAAGCTCATATGTGGCTCTCAGTGA 240
 DB 181 GGCCTGCTTCAAGTCAAGCGCTCCAGAGGCTGGAAGCTCATATGTGGCTCTCAGTGA 240

DB 181 GGCCTGCTTCAAGTCAAGCGCTCCAGAGGCTGGAAGCTCATATGTGGCTCTCAGTGA 240
 QY 241 CATGTGTGCTCTAAGCCTCAGGCCCATGAGCCCATCATCATATGACCGCTTCACTC 300
 DB 241 CATGTGTGCTCTAAGCCTCAGGCCCATGAGCCCATCATCATATGACCGCTTCACTC 300
 QY 301 TCAGAGGTACGACGAGGCGGGAACCTTCACTGCGAGATGATCATCCAAATGTGAGCC 360
 DB 301 TCAGAGGTACGACGAGGCGGGAACCTTCACTGCGAGATGATCATCCAAATGTGAGCC 360
 QY 361 CAGTGATTCGGGGAGACATCATGATGACGCTCCAGAACAGTCCGCTCATGATCTGCTTA 420
 DB 361 CAGTGATTCGGGGAGACATCATGATGACGCTCCAGAACAGTCCGCTCATGATCTGCTTA 420
 QY 421 CCTTACCGTCCAAAGTATATGGAGAGCTGTTCATTTCCAGTGTAACTTTGATGTCGTA 480
 DB 421 CCTTACCGTCCAAAGTATATGGAGAGCTGTTCATTTCCAGTGTAACTTTGATGTCGTA 480
 QY 481 GAATGAACCTTGTGAGATTAATGTTACCTCCACACCTGAGCCGCGGATATATTTTC 540
 DB 481 GAATGAACCTTGTGAGATTAATGTTACCTCCACACCTGAGCCGCGGATATATTTTC 540
 QY 541 CTGGAGCTCGGTCTCTCTGTCAGCCATTCAAGCTATATTTGTTCCGAGCCAGCGA 600
 DB 541 CTGGAGCTCGGTCTCTCTGTCAGCCATTCAAGCTATATTTGTTCCGAGCCAGCGA 600
 QY 601 CCTTCAAGGCTGAGATGATGATGCTGCTGACCCACAGACATGAGACTTTGACTTG 660
 DB 601 CCTTCAAGGCTGAGATGATGATGCTGCTGACCCACAGACATGAGACTTTGACTTG 660
 QY 661 CGTGAGCTGCTGAGAGAGCCCTGAGGCCCGCAAGTCTGCACTGTAATCTCATGTGAT 720
 DB 661 CGTGAGCTGCTGAGAGAGCCCTGAGGCCCGCAAGTCTGCACTGTAATCTCATGTGAT 720
 QY 721 TCGGTCTCCCAAGAGACAGAGGAGGTGATTAATTCAGGTGATATATCAAGTTTACC 780
 DB 721 TCGGTCTCCCAAGAGAGAGGAGGTGATTAATTCAGGTGATATATCAAGTTTACC 780
 QY 781 GAGTTAGGTTTCTTCAATTCCTACTTGGGGCAAGTTGGACTGTGAGCAGCACCAT 840
 DB 781 GAGTTAGGTTTCTTCAATTCCTACTTGGGGCAAGTTGGACTGTGAGCAGCACCAT 840
 QY 841 GCTTTCAGAGCGGAGGTGTACTTCAATACGCTGTCTGCTCCGCCGCTGCTTGTG 900
 DB 841 GCTTTCAGAGCGGAGGTGTACTTCAATACGCTGTCTGCTCCGCCGCTGCTTGTG 900
 QY 901 TGGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 TGGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 TCAATTTCAAAAGAAATCTGAAAAGAGAGCAAAAGAAAGAAAGAAAGAAAGAAAG 1020
 DB 961 TCAATTTCAAAAGAAATCTGAAAAGAGAGCAAAAGAAAGAAAGAAAGAAAGAAAG 1020
 QY 1021 AATGAAAACCTCGGCTCAATTCAGATGAAACAAAGACACAGAAACCGCTTCTCTCC 1080
 DB 1021 AATGAAAACCTCGGCTCAATTCAGATGAAACAAAGACACAGAAACCGCTTCTCTCC 1080
 QY 1081 TCCCAAAATCTTGAATCTGATCAGTATCTGAAACAAAGAAAGAAAGAAAGAAAGAAAG 1140
 DB 1081 TCCCAAAATCTTGAATCTGATCAGTATCTGAAACAAAGAAAGAAAGAAAGAAAGAAAG 1140
 QY 1141 CCAGGCGGCTGATCAAGCTCCACCCAGGCGAGCAAGTATCATCAGAGCTTTTAACT 1200
 DB 1141 CCAGGCGGCTGATCAAGCTCCACCCAGGCGAGCAAGTATCATCAGAGCTTTTAACT 1200
 QY 1201 GGCAGTCTCTGAGAGTCAAGTATCAATCAAGTATGATAG 1240
 DB 1201 GGCAGTCTCTGAGAGTCAAGTATCAATCAAGTATGATAG 1240

RESULT 2
 ABK13028

QY 1 AGGTGTGAGTCCACCAACAGTGTGATCAATTCTAGGCTGCCATTAACAAGCACCAT 60
|||||
Db 159444 AGGTGTGAGTCCACCAACAGTGTGATCAATTCTAGGCTGCCATTAACAAGCACCAT 159503
|||||
QY 61 AACCTGGTGGCTTAGAACATGAAAGCATTTGCTCAGCGTTCCAGAAAGCTGTAGGTT 119
|||||
Db 159504 AACCTGGTGGCTTAGAACATGAAAGCATTTGCTCAGCGTTCCAGAAAGCTGTAGGTT 159562
|||||

Search completed: April 29, 2003, 14:51:53
Job time : 4716.78 secs

/note="GC score = 10.60 (249bp); Region: GC content"
 /evidence=not_experimental
 repeat_region 15369.16268
 /rpt_family="MERVL"
 complement(16623.16714)
 /note="GRAIL score = 72.000%, comment = good"
 /evidence=not_experimental
 exon 16661.16730
 /note="MZF score = 93.58"
 /evidence=not_experimental
 repeat_region 16880.17318
 /evidence=not_experimental
 /rpt_family="MLTIC"
 repeat_region 17326.17436
 /evidence=not_experimental
 /rpt_family="LIR16C"
 complement(17540.17675)
 /evidence=not_experimental
 /rpt_family="MIR"
 complement(17805.17864)
 /evidence=not_experimental
 /rpt_family="MLTIC"
 exon 18076.18098
 /note="GRAIL score = 45.000%, comment = marginal shadow"
 /evidence=not_experimental
 complement(18219.18349)
 /note="GenScan score = 2.998, comment = Internal-exon 131
 bp frame: 1 phase: 2"
 /evidence=not_experimental
 repeat_region complement(18717.19211)
 /evidence=not_experimental
 /rpt_family="MLTIC"
 exon complement(19252.19399)
 /note="GRAIL score = 71.000%, comment = good"
 /evidence=not_experimental
 exon 19269.19350
 /note="GRAIL score = 81.000%, comment = excellent shadow"
 /evidence=not_experimental
 complement(19379.19399)
 /note="Xpound exon prediction, score = 60% (0%)"
 /evidence=not_experimental
 exon complement(19501.19881)
 /note="MZF score = 50.6%"
 /evidence=not_experimental
 repeat_region 19945.20020
 /note="homology = 100.00%, score = 38, counts = 2"
 /evidence=not_experimental
 /rpt_type=tandem
 repeat_region /rpt_unit=ttcttcacactctctgtgagagtcggttatcgc
 20223.20643
 /evidence=not_experimental
 /rpt_family="MLTIF"
 repeat_region 20645.20910
 /evidence=not_experimental
 /rpt_family="MLTIF"
 repeat_region 20961.21007
 /evidence=not_experimental
 /rpt_family="MLTIF"
 repeat_region 21483.21781
 /evidence=not_experimental
 /rpt_family="ALUB"
 complement(21504.21756)
 /note="MZF score = 79.4%"
 /evidence=not_experimental
 complement(22325.22806)
 /evidence=not_experimental
 /rpt_family="MLTIC"

Query Match 9.6%; Score 119; DB 9; Length 40205;
 Best Local Similarity 100.0%; Pred. No. 17e-17;
 Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGTGTGAGTCCAGCACAAGTGTGATCACTTCTAGCTGCCATACAAACACCAT 60
 |||||
 Db 18694 AGGTGTGAGTCCAGCACAAGTGTGATCACTTCTAGCTGCCATACAAACACCAT 18753
 Oy 61 AACCTGTGGCTTAGAACATGGAAGGCAATTTGCTCAGGTTCCAGAACTGTAGCTT 119
 |||||
 Db 18754 AACCTGTGGCTTAGAACATGGAAGGCAATTTGCTCAGGTTCCAGAACTGTAGCTT 18812

RESULT 15
 AF064860
 LOCUS
 DEFINITION
 Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
 sequence.
 AF064860
 VERSION
 AF064860.2 GI:18958624
 KEYWORDS
 HTG; HTGS; DRAFT.
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 170121)
 Hattori, M., Fujiyama, A., Taylor, T.D., Matanabe, H., Yada, T.,
 Park, H.-S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D.-K., Soeda, E.,
 Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K.,
 Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R.,
 Patterson, D., Reichwald, K., Rump, A., Schillabel, M.B., Schudy, A.,
 Zimmermann, W., Rosenthal, A., Kudoh, J., Kawasaki, K., Asakawa, S.,
 Shintani, A., Sasaki, T., Nagamine, K., Mitsuyama, S.,
 Antonarakis, S.E., Minoshima, S., Shimizu, N., Nordiek, G.,
 Honisch, K., Brandt, P., Scharfe, M., Scheen, O., Desario, A.,
 Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S.,
 Hennig, S., Riesselmann, L., Dagand, E., Haaf, T., Wehmeyer, S.,
 Borzym, K., Gardiner, K., Nizetic, D., Francis, F., Lehnach, H.,
 Reinhardt, R. and Yaspo, M. Laure.
 The DNA sequence of human chromosome 21
 Nature 405 (6784), 311-319 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20289799
 10830953
 REFERENCE
 AUTHORS
 Rump, A., Dagand, E., Hildmann, T., Nordiek, G., Drescher, B.,
 Weber, J., Schattevoy, R., Yaspo, M.-L. and Rosenthal, A.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
 REFERENCE
 AUTHORS
 Rump, A., Dagand, E., Hildmann, T., Nordiek, G., Drescher, B.,
 Weber, J., Schattevoy, R., Yaspo, M.-L., Rosenthal, A., Yaspo, M.-L. and
 Rosenthal, A.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
 REFERENCE
 AUTHORS
 Rump, A., Dagand, E., Hildmann, T., Nordiek, G., Drescher, B.,
 Weber, J., Schattevoy, R., Yaspo, M.-L. and Rosenthal, A.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
 COMMENT
 On Feb 27, 2002 this sequence version replaced gi:3171153.
 FEATURES
 source
 1. 170121
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 /clone="PAC 70124"

BASE COUNT 45184 a 36756 c 37509 g 50672 t
 ORIGIN

Query Match 9.6%; Score 119; DB 9; Length 170121;
 Best Local Similarity 100.0%; Pred. No. 1.8e-17;
 Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contiguous gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

source
1. .175861
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-147E11"
/clone_id="RPC1-23 Female Mouse BAC"
1. .28835
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
28936. .29562
/note="assembly_fragment"
29663. .30328
/note="assembly_fragment"
30429. .32005
/note="assembly_fragment"
32106. .33128
/note="assembly_fragment"
33229. .34512
/note="assembly_fragment"
34613. .36079
/note="assembly_fragment"
36180. .37483
/note="assembly_fragment"
37584. .39166
/note="assembly_fragment"
39267. .41065
/note="assembly_fragment"
41166. .43203
/note="assembly_fragment"
43304. .47489
/note="assembly_fragment"
47590. .50604
/note="assembly_fragment"
50705. .56370
/note="assembly_fragment"
56471. .64079
/note="assembly_fragment"
64180. .72619
/note="assembly_fragment"
72720. .81566
/note="assembly_fragment"
81667. .96774
/note="assembly_fragment"
96875. .119647
/note="assembly_fragment"
119748. .147772
/note="assembly_fragment"
147873. .174557
/note="assembly_fragment"
174658. .175861
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 48903 a 39540 c 39029 g 46286 t 2103 others
ORIGIN
Query Match 14.2% Score 176.4; DB 2; Length 175861;
Best Local Similarity 73.3%; Pred.No.4.2e-31;
Matches 239; Conservative 0; Mismatches 86; Indels 1; Gaps 1;
QY 112 TGAAGTCTGGGTGATGAAGTATGAGAGGCCGCCAGATGACACAGTCTCTGAA 171
Db 120532 TCAAGCTCGGATTCAGATTACATGACATGATGAAGGTCCTGAGATGTATACAGTCTTAA 120473
QY 172 GGGCTCCAGAGCTCGCTTAACTGACCGT-CTCCAGGGCTGGAAGCTATATGTGGG 230
Db 120472 GGACTCAGAGGCTCACTTAACTGACCGTGGGCTCAGGCTGGAACTTCTATGTGA 120413
QY 231 CTCTCAGTGCATGGGTGATGAAGTATGAGAGGCCGCCAGTATGACCATGACCATGAC 290
Db 120412 CTCTTACCAATATGGTGGTCTGATGATCTACACCAACCAAGACCATATATACCAACAC 120353
QY 291 GCTTCAACCTTCAGAGGTTCAGACCAAGGCGGGAACCTTCACTCGAGATGATATCCACA 350

```


RESULT 10	AX380408	754 bp	DNA	linear	PAT 18-MAR-2002
LOCUS	AX380408				
DEFINITION	Sequence 13 from Patent WO0200710.				
ACCESSION	AX380408				
VERSION	AX380408.1	GI:19575338			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	1. Welcher, A.A., Sarmiento, U.M., Schuit, H.J. and Chute, H.T.				
AUTHORS	77-like molecules and uses thereof				
TITLE	Patent: WO 0200710-A 13 03-JAN-2002.				
JOURNAL	Amgen, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..754				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	53..724				
CDS	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD28222.1"				
	/db_xref="GI:19575339"				
	/translation="MEGSRDVLAVILIAQLTASGSSYOIEGPONVTLKDEAHF				
	NCVYTHGKRLMTNLMVYLSLTGQPIITNNRFTYASNSDSEFISLLIHVQPS				
	DSGSVOCISLNSHGFGSAPLFSVQESTQNEIRTSANMRTKADPEIKSGKENYGS				
	SDPKAAQVTSLEPPKSAEVSLEPERSSLEFYQELNKHQHPATHPRVPIASQKRV				
	NVTLV"				
BASE COUNT	220 a 191 c 175 g 168 t				
ORIGIN					
Query Match	15.0%: Score 185.6; DB: 6; Length 754;				
Best Local Similarity	73.8%: Pred. No. 2.2e-33;				
Matches 236; Conservative	0; Mismatches 84; Indels 0; Gaps 0;				
QY	115 AGTTCTGGGCTCTGGTAATGAGTCATGAAAGGCCCCAGAAATGACAGCTCTGAAAGG	174			
Db	109 AGCTTCGGGATCCAGTTATAGATCATGAAAGGTCCTCAGAAATGTAACAGCTCTAAAGA	168			
QY	175 CTCGCCAGGCTGCTTCAACTGCACCGCTCTCCAGGGCTGGAAGCTCATCATGTGGGCT	234			
Db	169 CTCAGAGGCTCACTCAACTGCACCGTACGCTACGCGCTGGAAGCTTCTCATGTGGACTCT	228			
QY	235 CAGTGCATGTGNGGCTTAACGTCAGGCCCCATGGAGGCCATCATCACAACCGCTT	294			
Db	229 TAACCAATAGTGTGCTGATGCTCAGCTCACCACCAACAGACCATCATCACAACACGCTT	288			
QY	295 CACCTCTGAGAGTAGACGACGAGGCGGGAAGCTTCACTCGTGGAGATGATCATCACAATGT	354			
Db	289 CACCTATGCCAGTTTACACAGACACTGACAGAGTCTTCATCTCGGGTGTGATCATCATGATGT	348			
QY	355 GGAGGCCGATATTGCGGGGAACATGATGGAAGCCCTCAGAAAGCACTGCGCTGCATGATGC	414			
Db	349 GGAGCCGATGACTCGGATCCGTGTCATGACGCTTGACAGAACACCATGGGTTTGATGC	408			
QY	415 TGCTTACCTTACCGTCAAG	434			
Db	409 TGCTTCTCTCTCAGTGAAG	428			
RESULT 11					
LOCUS	AX380406	895 bp	DNA	linear	PAT 18-MAR-2002
DEFINITION	Sequence 11 from Patent WO0200710.				
ACCESSION	AX380406				
VERSION	AX380406.1	GI:19575336			
KEYWORDS					
ORGANISM	house mouse.				
SOURCE	Mus musculus				

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
B7-like molecules and uses thereof
Patent: WO 0200710-A 11 03-JAN-2002;

FEATURES
source
Location/Qualifiers
1..895
/organism="Mus musculus"
/db_xref="taxon:10090"
53..865
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28221.1"
/db_xref="GI:19575337"
/translation="MEGSRWDVLAVALVILIAQLTASGSSQIIIEGPONVTLKSEAF
NCVHTVSKLIMTINQMYSLITGSPITTNPFYASVNSDFSELIDVOPS
DSGSVQCSLONSHRGFGAFSLSYDSDISEBPALPTMAITLLAAVSLILLIIVLII
FCCCAARRKEESTYONETKRSAMNTINADEETKSKRENYGVSSDRAKAQTNLS
LPPSAVSLSPEKRSSSLPYOELNKHQPGATHPVSDIASPOKRVNVLV"

BASE COUNT 250 a 223 c 212 g 210 t

ORIGIN

Query Match 15.0%; Score 185.6; DB: 6; Length 895;
Best Local Similarity 73.8%; Pred. No. 2.2e-33;
Matches 236; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

OY 115 AGGTTCTGGGCTGGTATGAAGTCATGTAAGAGGCCCCAGAAATGCACAGTCTCTAAGGG 174
|||||
DB 109 AGCTTCGGGATCCATTTATCGATCATGTGAAGGTCCTCAGATGTMAACAGTCTTAAGGA 168
175 CTCGACGAGCTCGCTTCACACTGCACCGTCTCCAGGGCTGGAAGCTCATCATGTGGGCTCT 234
|||||
DB 169 CTCAGAGGCTCACTTCACACTGCACCGTGCATCAGCGGTGGAAGCTTCTCATGTGGACTCT 228
OY 235 CAGTACATGTGTGCTGAAGCGTGAAGCCCATGAGGCCCATCATCACCATTGAGCCGCTT 294
DB 229 TAACCAATATGGTGGTCTGAGAGTCTACCAACCAAGGACCATCATCACCAACACCGCTT 288
OY 295 CACCTTCAGAGAGTACGACGACAGGGGGGAGACCTTCAGCTCGGAGATGATCATCCACATGT 354
|||||
DB 289 CACCTATGCCAGTTTACAAACAGCAGTACAGAGCTTCACTCTCGGAGTTGATCATCCAGATGT 348
OY 355 GGAGCCCAAGTATTCGGGGAACATCAGATGACAGCCCTCGAGAAACAGTGCCTGCATGATG 414
|||||
DB 349 GCAGCCCAAGTACGTGGGATCCGTCAATGACGCTCGAGAAACAGCAGTGGTTTGATC 408
OY 415 TGCTTACCTTACCGTCCAAG 434
|||||
DB 409 TGCTTCTCTCTCAGTGAAG 428

RESULT 12
LOCUS BC004806 1556 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, RIKEN cDNA 201003020 gene, clone MGC:7960
IMAGE:3584645, mRNA, complete cds.
ACCESSION BC004806
VERSION BC004806.1 GI:13435932
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Strausberg, R.
Direct Submission
Submitted (21-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

LOCUS AC020851 196900 bp DNA linear HTG 15-JUL-2000
 DEFINITION Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33
 unorderd pieces.
 AC020851
 VERSION AC020851.2 GI:9211212
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Mouse
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 196900)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jul 15, 2000 this sequence version replaced gi:6686457.
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 1437240
 Center clone name: RP21-467L12

Summary Statistics
 Consensus quality: 164618 bases at least Q40
 Consensus quality: 180887 bases at least Q30
 Consensus quality: 183777 bases at least Q20
 Estimated insert size: 14700; pulse field gel estimation
 Estimated insert size: 193700; sum-of-contigs estimation
 Quality coverage: 5.67 in Q20 bases; pulse field gel estimation
 Quality coverage: 4.3 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      1541: contig of 1541 bp in length
*      1542      1641: gap of unknown length
*      1642      3492: contig of 1851 bp in length
*      3493      3593: gap of unknown length
*      3593      4738: contig of 1146 bp in length
*      4739      4838: gap of unknown length
*      4839      7248: contig of 2410 bp in length
*      7249      7348: gap of unknown length
*      7349      9178: contig of 1830 bp in length
*      9179      9278: gap of unknown length
*      9279      11493: contig of 2215 bp in length
*      11494      11593: gap of unknown length
*      11594      13607: contig of 2014 bp in length
*      13608      13707: gap of unknown length
*      13708      15946: contig of 2239 bp in length
*      15947      16046: gap of unknown length
*      16047      18392: contig of 2346 bp in length
*      18393      18492: gap of unknown length
*      18493      21231: contig of 2739 bp in length
*      21232      21331: gap of unknown length
*      21332      24662: contig of 3331 bp in length
*      24663      24762: gap of unknown length
*      24763      27140: contig of 2378 bp in length
*      27141      27240: gap of unknown length
*      27241      29950: contig of 2710 bp in length
*      29951      30050: gap of unknown length
*      30051      32245: contig of 2195 bp in length
*      32246      32345: gap of unknown length

```

```

*      32346      35688: contig of 3343 bp in length
*      35689      35788: gap of unknown length
*      35789      40387: contig of 4599 bp in length
*      40388      40487: gap of unknown length
*      40488      43580: contig of 3093 bp in length
*      43581      43680: gap of unknown length
*      43681      46437: contig of 2757 bp in length
*      46438      46537: gap of unknown length
*      46538      50184: contig of 3647 bp in length
*      50185      50284: gap of unknown length
*      50285      54572: contig of 4288 bp in length
*      54573      54672: gap of unknown length
*      54673      60227: contig of 5555 bp in length
*      60228      60327: gap of unknown length
*      60328      68024: contig of 7697 bp in length
*      68025      68124: gap of unknown length
*      68125      78610: contig of 10486 bp in length
*      78611      78710: gap of unknown length
*      78711      88151: contig of 9441 bp in length
*      88152      88251: gap of unknown length
*      88252      96257: contig of 8006 bp in length
*      96258      96357: gap of unknown length
*      96358      106227: contig of 9870 bp in length
*      106228      106327: gap of unknown length
*      106328      115795: contig of 9468 bp in length
*      115796      115895: gap of unknown length
*      115896      126363: contig of 10468 bp in length
*      126364      126463: gap of unknown length
*      126464      137189: contig of 10726 bp in length
*      137190      137289: gap of unknown length
*      137290      146378: contig of 9089 bp in length
*      146379      146478: gap of unknown length
*      146479      161686: contig of 15208 bp in length
*      161687      161786: gap of unknown length
*      161787      176149: contig of 14363 bp in length
*      176150      176250: gap of unknown length
*      176250      196900: contig of 20651 bp in length.

```

FEATURES
 source
 1..196900
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP21-467L12"
 /clone_1bp="RP21 mouse PAC library 21"

BASE COUNT 51456 a 43823 c 44608 g 53666 t 3347 others
 ORIGIN

Query Match 15.1%; Score 187.4; DB 2; Length 196900;
 Best Local Similarity 73.5%; Pred. No. 1e-33;
 Matches 239; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```

QY 112 TGTAGTTCTGGTCTGTAATGAATGATGAAGCCGCCAGAAATGCAAGTCTGAA 171
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67442 TCTAGCTTCCCGATCCAGTATCATGATCAGATGATGATGATGATGATGATG 67501

QY 172 GGGCTCCAGGCTGCTCACTCACTGACGCTCCAGGCTGGAAGCTCATCATGTGGG 231
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67502 GGACTCAGAGCTCACTTCACTGACGCTGACCTGACCTGAGAGCTTCTCATGTGAC 67561

QY 232 TGTGAGTACATGTGTGTGCTAAGCTCAGGCCCATGAGGCCATCATCAACATGACCG 291
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67562 TCTTAACCAATATGAGTGTGAGTGTGACCAACCAAGAGCCATCATCAACCAACCG 67621

QY 292 CTTACACCTCTAGAGGTACGACGAGGCGGGAACCTTCACTCTCGAGATGATCATCCACA 351
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67622 CTTACACCTATGCCAGTTTCAACACGACCTGACAGGTTCTTCGAGATGATCATTCATGA 67681

QY 352 TGTGAGCCCGATGATTCGAGGAGCATCATGATGAGAGGAGTGCCTGATGAG 411
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67682 TGTGAGCCCGATGATTCGAGGAGCATCATGATGAGAGGAGTGCCTGATGAG 67741

QY 412 ATCTGCTTACCTTACCGTCAAGTT 436
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67742 ATCTGCTTACCTTACCGTCAAGTT 67766

```



```

repeat_region      /rpt_family="MER34"
                    complement(17633..17969)
                    /evidence=not_experimental
                    /rpt_family="AluSg1"
repeat_region      /evidence=not_experimental
                    complement(18004..18334)
                    /rpt_family="L1m17"
repeat_region      /evidence=not_experimental
                    complement(18335..18632)
                    /evidence=not_experimental
                    /rpt_family="HERV17"
exon                /note="GRAIL, score = 51.00%, comment = good shadow"
                    /evidence=not_experimental
exon                /note="GRAIL, score = 66.00%, comment = good"
                    /evidence=not_experimental
                    complement(18643..24647)
repeat_region      /evidence=not_experimental
                    /rpt_family="HERV17"
                    complement(19228..19361)
exon                /note="MZF, score = 92.8%"
                    /evidence=not_experimental
                    20900..21135
exon                /note="MZF, score = 51.4%"
                    /evidence=not_experimental
                    21150..21275
exon                /note="Xpound exon prediction, score = 74% (0%)"
                    /evidence=not_experimental

```

```

Query Match      25 7%; Score 318.6; DB 9; Length 142742;
Best Local Similarity 98.8%; Pred. No. 7e-65;
Matches 321; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 115 AGGTTCTGGGCTCTGTAATGATGATAGAGAGGCCCCAGAAATGACAGTCTCGAAGG 174
    |||||||
Db 44273 AGGTTCTGGGCTCTGTAATGATGATAGAGAGGCCCCAGAAATGACAGTCTCGAAGG 44332
    |||||||
QY 175 CTCCAGGCTGCTTCACTGACGCTGCCAGGGCTGGAAGCTCATCATGTGGCTCT 234
    |||||||
Db 44333 CTCCAGGCTGCTTCACTGACGCTGCCAGGGCTGGAAGCTCATCATGTGGCTCT 44392
    |||||||
QY 235 CAGTGACATGATGCTGTAAGCGTCAGGCGGAGGCGGAGCCCATATCATCAATGACCGCTT 294
    |||||||
Db 44393 CAGTGACATGATGCTGTAAGCGTCAGGCGGAGGCGGAGCCCATATCATCAATGACCGCTT 44452
    |||||||
QY 295 CACCTCTCAGAGGTACAGCAGGCGGAGGAACTTCACTCGAGATGATCATCCCAATGT 354
    |||||||
Db 44453 CACCTCTCAGAGGTACAGCAGGCGGAGGAACTTCACTCGAGATGATCATCCCAATGT 44512
    |||||||
QY 355 GGAGCCCATGATTCGGGGAAATCATGATCAGCCTCCAGAACAGTGCCTGCATGATC 414
    |||||||
Db 44513 GGAGCCCATGATTCGGGGAAATCATGATCAGCCTCCAGAACAGTGCCTGCATGATC 44572
    |||||||
QY 415 TGCTTACCTTACGCTCAAGTATG 439
    |||||||
Db 44573 TGCTTACCTTACGCTCAAGTATG 44597
    |||||||

```

```

RESULT 8
LOCUS      HS21C080      340000 bp      DNA      linear      PRI 24-MAY-2000
DEFINITION Homo sapiens chromosome 21 segment HS21C080.
ACCESSION AL163280 AF001735 BA000005
VERSION    AL163280.2 GI:7717369
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 340000)
AUTHORS   Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Bleichschmidt,K.,

```

TITLE JOURNAL

COMMENT

Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
 Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
 Zimmermann,W., Rosenthal,A., Kudo,J., Shibuya,K., Kawasaki,K.,
 Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,
 Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G.,
 Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
 Reichelt,J., Kauer,G., Bioecker,H., Ramser,J., Beck,A., Klages,S.,
 Hennig,S., Riesselmann,L., Dagand,E., Wehmeyer,S., Borzym,K.,
 Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
 Yaspo,M.L.

Direct Submission

Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
 Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
 Group * Institute of Molecular Biotechnology, Genome Analysis *
 Keio University School of Medicine, Dept. of Molecular Biology *
 GSF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
 Genetics (addresses see below)

The Chromosome 21 Mapping and Sequencing Consortium consists of
 * RIKEN Genomic Sciences Center, Human Genome Research Group, *
 Sagamihara 228-8555, Japan,
 * e.mail: sakaki@gsf.riken.go.jp/
 * URL: http://hgp.gsc.riken.go.jp/
 and
 * Institute of Molecular Biotechnology, Genome Analysis, *
 Beutenbergstrasse 11, D-07745 Jena, Germany,
 * e.mail: gscj-submit@genome.imb-jena.de
 * URL: http://genome.imb-jena.de/
 and
 * Keio University School of Medicine, Dept. of Molecular Biology, *
 Tokyo 160-8582, Japan,
 * e.mail: shimizu@mb.med.keio.ac.jp/
 * URL: http://adenine.dmb.med.keio.ac.jp/
 and
 * GSF, Dept. of Genome Analysis,
 Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
 info.genome@gsf.de
 * URL: http://genome.gbf.de/
 and
 * Max-Planck Institute for Molecular Genetics,
 Innestrasse 73, D-14195 Berlin, Germany,
 * e.mail: info-chr21@molgen.mpg.de
 * URL: http://chr21.rz-berlin.mpg.de/
 Location/Qualifiers

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1..125946
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"

```

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P16019, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P70124"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000

```

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000

```

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000

```

source

```

repeat_region complement(140. .415)
/rpt_family="MLT2E"
/evidence-not_experimental
repeat_region complement(68. .881)
/evidence-not_experimental
/rpt_family="MER58A"
exon complement(1078. .1162)
/note="GRAIL, score = 86.000%, comment = excellent shadow"
/evidence-not_experimental
repeat_region 1114. .1403
/evidence-not_experimental
/rpt_family="MLT1I"
repeat_region 1481. .1603
/evidence-not_experimental
/rpt_family="L2"
exon 1576. .1710
/note="MZEf, score = 66.1%"
/evidence-not_experimental
1576. .1710
/note="GRAIL, score = 99.000%, comment = excellent"
/evidence-not_experimental
repeat_region complement(1693. .1741)
/evidence-not_experimental
/rpt_family="MIR"
repeat_region complement(1780. .2326)
/evidence-not_experimental
/rpt_family="MLT1E"
exon complement(1834. .1955)
/note="GRAIL, score = 50.000%, comment = good shadow"
/evidence-not_experimental
2629. .2831
/evidence-not_experimental
repeat_region /rpt_family="AlusX"
3012. .3122
/evidence-not_experimental
repeat_region /evidence-not_experimental
/rpt_family="MLT1J"
3123. .3478
/evidence-not_experimental
/rpt_family="THE1B"
repeat_region 3479. .3618
/evidence-not_experimental
/rpt_family="MLT1J"
exon complement(3502. .3540)
/note="GRAIL, score = 65.000%, comment = good shadow"
/evidence-not_experimental
4674. .4781
/note="GRAIL, score = 42.000%, comment = marginal"
/evidence-not_experimental
exon 4712. .4749
/evidence-not_experimental
/note="Xpound exon prediction, score = 70% (0%)"
repeat_region 5385. .6104
/evidence-not_experimental
/rpt_family="HERV1"
exon 6682. .6751
/note="MZEf, score = 95.1%"
/evidence-not_experimental
6746. .6751
/note="Genscan, score = 1.83%, comment = Initial_exon 6 bp.
frame: 1 phase: 0"
repeat_region /evidence-not_experimental
6901. .7339
/evidence-not_experimental
/rpt_family="MLT1C"
repeat_region 7347. .7457
/evidence-not_experimental
/rpt_family="L1P6"
repeat_region complement(7561. .7696)
/evidence-not_experimental
/rpt_family="MIR"
repeat_region complement(7819. .8220)
/evidence-not_experimental
/rpt_family="MLT1J"

```

```

exon 8097. .8119
/note="GRAIL, score = 45.000%, comment = marginal"
/evidence-not_experimental
repeat_region complement(8738. .9232)
/evidence-not_experimental
/rpt_family="MLT1C"
exon complement(9273. .9420)
/note="GRAIL, score = 71.000%, comment = good"
/evidence-not_experimental
exon 9290. .9371
/note="GRAIL, score = 81.000%, comment = excellent shadow"
/evidence-not_experimental
repeat_region 9966. .10041
/note="homology = 100.00%, score = 38, counts = 2"
/evidence-not_experimental
/rpt_type=tandem
/rpt_family="MLT1E"
repeat_region 10244. .10664
/evidence-not_experimental
/rpt_family="MLT1E"
repeat_region 10665. .11028
/evidence-not_experimental
/rpt_family="L1MC/D"
repeat_region 11318. .11417
/evidence-not_experimental
/rpt_family="L1MC2"
repeat_region 11503. .11818
/evidence-not_experimental
/rpt_family="Alus9"
exon complement(11525. .11777)
/note="MZEf, score = 79.4%"
/evidence-not_experimental
repeat_region complement(12346. .12827)
/evidence-not_experimental
/rpt_family="MLT1C"
repeat_region 12902. .13180
/evidence-not_experimental
/rpt_family="Alus9"
exon complement(13057. .13161)
/note="MZEf, score = 76.1%"
/evidence-not_experimental
repeat_region complement(13226. .13364)
/evidence-not_experimental
/rpt_family="L1MC2"
repeat_region complement(13613. .13908)
/evidence-not_experimental
/rpt_family="AlusX"
repeat_region complement(14317. .14614)
/evidence-not_experimental
/rpt_family="L1P6"
exon complement(14962. .15036)
/note="MZEf, score = 93%"
/evidence-not_experimental
repeat_region 15255. .15299
/evidence-not_experimental
/rpt_family="L1P6"
exon complement(15469. .15497)
/note="Xpound exon prediction, score = 62% (0%)"
/evidence-not_experimental
repeat_region 16045. .16768
/evidence-not_experimental
/rpt_family="MER21A"
exon complement(16555. .16655)
/note="GRAIL, score = 54.000%, comment = good"
/evidence-not_experimental
repeat_region 16823. .16850
/evidence-not_experimental
/rpt_family="MER34"
repeat_region complement(16851. .17208)
/evidence-not_experimental
/rpt_family="THE1A"
repeat_region 17209. .17632
/evidence-not_experimental

```



```

Db 694 ATATCAAGTTTACCGAGTTAGTTTTCATTCCTACTTGGGGCAAGTTGGACTTGG 753
Qy 826 ACTGACAGGACCATGCTTCTGACGGCGAGCTGACTTCAATACGCTGCTGCTG 885
Db 754 ACTGACAGGACCATGCTTCTGACGGCGAGCTGACTTCAATACGCTGCTGCTG 813
Qy 886 CCGCGCTGCTTGTGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
Db 814 CCGCGCTGCTTGTGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873
Qy 946 AAGAGATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGAGAGAGAGAAC 1005
Db 874 AAGAGATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGAGAGAGAGAAC 933
Qy 1006 TGAACAGAGAAAGTGAAGAAATGAAAACCTCCGCTACATTCAGATGACAAAAGACACAGA 1065
Db 934 TGAACAGAGAAAGTGAAGAAATGAAAACCTCCGCTACATTCAGATGACAAAAGACACAGA 993
Qy 1066 AACGCTTCTCTCCCTCCCAATCTGTGAATCCAGATCCGAGAACAAAGAAACAGTAG 1125
Db 994 CACGCTTCTCTCCCTCCCAATCTGTGAATCCAGATCCGAGAACAAAGAAACAGTAG 1053
Qy 1126 CTGTGGCCCTCTCTCCACAGGGGCTGATCAACGTCACCCAGGCGACAGATCATCCACA 1185
Db 1054 CTGTGGCCCTCTCTCCACAGGGGCTGATCAACGTCACCCAGGCGACAGATCATCCACA 1113
Qy 1186 GGCCTCTTTTAATCTGGCCAGCTCTGAGAAAGCTAGTAATACACTGTAGTAG 1240
Db 1114 GGCCTCTTTTAATCTGGCCAGCTCTGAGAAAGCTAGTAATACACTGTAGTAG 1168

RESULT 5
AX380402 1139 bp DNA linear PAT 18-MAR-2002
LOCUS Sequence 7 from Patent WO0200710.
DEFINITION AX380402
ACCESSION AX380402
VERSION AX380402.1 GI:19575332
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-1 like molecules and uses thereof.
JOURNAL Patent: WO 0200710-A 7 03-JAN-2002;
Amgen, Inc. (US)
FEATURES
source
1. 1139
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 1134
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28219.1"
/db_xref="GI:19575333"
/translation="MYAGAMENDPPGSSGNVIEGPONAVIKSQARFNCYSOG
WKLIMALSDMVLVSRPEPITINDRTSQRIDGGLTSEKTIHNPESDGNTRC
SLONSRLHGSAYLTVOGELFIPSVNLVAENEPEDEVCLPSHWRPLDISMELGL
VSHSYFVPEPESDLOSAYSIATLPOSNGTULCVATWSLAKRSATVNLIVIRCPQ
DTGGGNGINGVSLPSLSESLPTGKVLGIGLAGLLPLPTGLTIRCCCGRRCCG
NCCGCCPCRRKRRGNLKRROTKLRKQVEKTEGATPTIOMNRPQTPILSLPVPNPV
ILNKETVAVALLTSLINVHPGQVTHRLLLIMPLKRSVLDL"
BASE COUNT 290 a 300 c 283 g 266 t
ORIGIN
Query Match 85.8%; Score 1064; DB 6; Length 1139;
Best Local Similarity 97.2%; Pred. No. 5.4e-242;
Matches 1102; Conservative 0; Mismatches 10; Indels 22; Gaps 1;
Qy 106 AGAAGCTGTGTGCTGCTGTGTAATGAAGTCAAGAGCCCGACAGATCAACAGT 165
111 1

```

```

Db 27 AGACCCACCCGGGTCTGGGCTGTGTAATGAAGTCAATAGAGGCCCCCAAAATGC AAGAGT 86
Qy 166 CCTGAAGGGCTCCCGAGGCTCGCTTCAACTGACACCGTCTCCCGAGGGCTGGAAGCTATCAT 225
Db 87 CCTGAAGGGCTCCCGAGGCTCGCTTCAACTGACACCGTCTCCCGAGGGCTGGAAGCTATCAT 146
Qy 226 GTGGGCTCTCACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 285
Db 147 GTGGGCTCTCACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 206
Qy 286 TGACCGCTTACCTCTCAGAGGTACGACAGGCGGGGAACTTACCTCGAGATGATCAT 345
Db 207 TGACCGCTTACCTCTCAGAGGTACGACAGGCGGGGAACTTACCTCGAGATGATCAT 266
Qy 346 CCACAAATGTGAGGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405
Db 267 CCACAAATGTGAGGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 326
Qy 406 GCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 465
Db 327 GCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 386
Qy 466 TCTTGTAGTCTGCTGAGAAATGAACCTTGTGAAGTACTTGTACCTCAGACTGAGACCG 525
Db 387 TCTTGTAGTCTGCTGAGAAATGAACCTTGTGAAGTACTTGTACCTCAGACTGAGACCG 446
Qy 526 GCTCCCGGATATTTCTCTGGGAGCTGGTCTCTGCTCAGCCATTACATATTATTGTT 585
Db 447 GCTCCCGGATATTTCTCTGGGAGCTGGTCTCTGCTCAGCCATTACATATTATTGTT 506
Qy 586 TCCGGAGCCCGAGGACCTTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 645
Db 507 TCCGGAGCCCGAGGACCTTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 566
Qy 646 TGGGACTTTGACTTGGCTGCTACCTGGAAGACCTGGAAGGCCCGCAAGTGTCAACTGT 705
Db 567 TGGGACTTTGACTTGGCTGCTACCTGGAAGACCTGGAAGGCCCGCAAGTGTCAACTGT 626
Qy 706 AATCTCACTGTGATTCGGTGTCCCAAGACCTGAGGTGATTAATATTCCAGGTGT 765
Db 627 AATCTCACTGTGATTCGGTGTCCCAAGACCTGAGGTGATTAATATTCCAGGTGT 686
Qy 766 ATTATCAATTTACCGAGATTTAGTTTTCATTTGCTTACCTTGGGCAAGTGTGACTGG 825
Db 687 ATTATCAATTTACCGAGATTTAGTTTTCATTTGCTTACCTTGGGCAAGTGTGACTGG 746
Qy 826 ACTAGCAGCACCAGTCTTCTACAGCCGACGTGTACTTACATATACGCTGCTGCTG 885
Db 747 ACTAGCAGCACCAGTCTTCTACAGCCGACGTGTACTTACATATACGCTGCTGCTG 806
Qy 886 CCGCGCTGCTGTGTGGCTGCACTGCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
Db 807 CCGCGCTGCTGTGTGGCTGCACTGCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 866
Qy 946 AAGAGATTTCTGATTCATTTCAAAAGAAATCTGAAAAGAGAGAGAGAGAGAGAAC 1005
Db 867 AAGAG-----GAAATCTGAAAAGAGAGAGAGAGAGAGAAC 904
Qy 1006 TGAACAGAGAAAGTGAAGAAATGAAAACCTCCGCTACATTCAGATGACAAAAGACACAGA 1065
Db 905 TGAACAGAGAAAGTGAAGAAATGAAAACCTCCGCTACATTCAGATGACAAAAGACACAGA 964
Qy 1066 AACGCTTCTCTCCCTCCCAATCTGTGAATCCAGATCCGAGAACAAAGAAACAGTAG 1125
Db 965 CACGCTTCTCTCCCTCCCAATCTGTGAATCCAGATCCGAGAACAAAGAAACAGTAG 1024
Qy 1126 CTGTGGCCCTCTCTCCACAGGGGCTGATCAACGTCACCCAGGCGACAGATCATCCACA 1185
Db 1025 CTGTGGCCCTCTCTCCACAGGGGCTGATCAACGTCACCCAGGCGACAGATCATCCACA 1084
Qy 1186 GGCCTCTTTTAATCTGGCCAGTCTGAGAAAGCTAGTAATACACTGTAGTAG 1239
Db 1085 GGCCTCTTTTAATCTGGCCAGTCTGAGAAAGCTAGTAATACACTGTAGTAG 1138

```


QY 356 GAGCCAGTATTTGGGGAACATGATGACGCTCCAGAAACAGTGGCTGCATGATCT 415
 DB 291 GAGCCAGTATTTGGGGAACATGATGACGCTCCAGAAACAGTGGCTGCATGATCT 350
 QY 416 GCTTACCTTACCGTCCAGTATGAGAGAGTGTTCATCCAGTAACTTGTACTC 475
 DB 351 GCTTACCTTACCGTCCAGTATGAGAGAGTGTTCATCCAGTAACTTGTACTC 410
 QY 476 GCTGAGATGAACCTTGTGAAGTACTGTCTACCTCAGTACAGTGGACCGGCTCCGGAT 535
 DB 411 GCTGAGATGAACCTTGTGAAGTACTGTCTACCTCAGTACAGTGGACCGGCTCCGGAT 470
 QY 536 ATTTCCTGGAGCTCGCTCTCTCTGTCAGCCATTCAAGCTATTTATTTGTTCCGAGCC 595
 DB 471 ATTTCCTGGAGCTCGCTCTCTCTGTCAGCCATTCAAGCTATTTATTTGTTCCGAGCC 530
 QY 596 AGCGACCTTCAAGTGCAGTGAAGATCTGCTGACCCAGCCAGAGAGATGGAGCTTTG 655
 DB 531 AGCGACCTTCAAGTGCAGTGAAGATCTGCTGACCCAGCCAGAGAGATGGAGCTTTG 590
 QY 656 ACTTGCCTGACCTGTGAAGAGCCGTGAAGCCCGCAAGTCTGCAACTGTAACTCACT 715
 DB 591 ACTTGCCTGACCTGTGAAGAGCCGTGAAGCCCGCAAGTCTGCAACTGTAACTCACT 650
 QY 716 GTGATTCGCTGCTCCCAAGACACTGAGAGTGTATTAATATTCAGGTGTATCAAGT 775
 DB 651 GTGATTCGCTGCTCCCAAGACACTGAGAGTGTATTAATATTCAGGTGTATCAAGT 710
 QY 776 TTACCGAGTTAGTTTATTTCTATGCTACTTGGGGCAAGTTGGACTGAGCTAGCAGC 835
 DB 711 TTACCGAGTTAGTTTATTTCTATGCTACTTGGGGCAAGTTGGACTGAGCTAGCAGC 770
 QY 836 ACCATGCTTCTGACGCGCAGCTGTACTTCAATACGCTGCTGCTGCGCGCTGCT 895
 DB 771 ACCATGCTTCTGACGCGCAGCTGTACTTCAATACGCTGCTGCTGCGCGCTGCT 830
 QY 896 TGTGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 DB 831 TGTGTGGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
 QY 956 CGTATTCATTTCAAAAGAAATCTGAAGAAAGAGAAAGCAAAAGAACTGAGACAGAA 1015
 DB 891 CGTATTCATTTCAAAAGAAATCTGAAGAAAGAGAAAGCAAAAGAACTGAGACAGAA 950
 QY 1016 AGTGAATATGAAGAACTCGGCTACATTCAGATGAACAAAGACACAGAAACCGCTTCT 1075
 DB 951 AGTGAATATGAAGAACTCGGCTACATTCAGATGAACAAAGACACAGAAACCGCTTCT 1010
 QY 1076 CTCCCTCCCAATCTGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1135
 DB 1011 CTCCCTCCCAATCTGTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1070
 QY 1136 CCTCAGCAGGCGGCTGATCAACGTCACCCAGGCGAGCAAGTCAATCCAGAGCTTCTTT 1195
 DB 1071 CCTCAGCAGGCGGCTGATCAACGTCACCCAGGCGAGCAAGTCAATCCAGAGCTTCTTT 1130
 QY 1196 AATTCGCGCAGTCTGTGAAGAGTCAATCAACTGTATATAG 1240
 DB 1131 AATTCGCGCAGTCTGTGAAGAGTCAATCAACTGTATATAG 1175

RESULT 4
 LOCUS AX380398 1168 bp DNA linear PAT 18-MAR-2002
 DEFINITION Sequence 3 from Patent WO0200710.
 ACCESSION AX380398
 VERSION AX380398.1 GI:19575328
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Welch, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
 TITLE B7-1 like molecules and uses thereof
 JOURNAL Patent: WO 0200710-A 3 03-JAN-2002;
 Amgen, Inc. (US)

FEATURES
 source Location/Qualifiers
 1..1168
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 8..1168
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD28217.1"
 /db_xref="GI:19575328"
 /translation="MVAGAMENRDPGSGSNEVIEGPQNAVRLKSGQARFNCVSOG
 WKILMALSDVIVLVSFPMPEIITNDFTSORYOGNFTSEMIHNHVPDSGNIC
 SLONSRLHSAVYLVOYMGELFIPSVNLVYVENEPCVTCIPSHWTLPIDSMELGIL
 VSHSYFVPEPSDLSAVSLATIPSNGLITLVATWKSIAKASVNLIVTRCO
 DTGGINIPGLSLPSLGLFPTMGVGLAGLMTLLTPCTLLTIRCCRRRCGC
 NCCRCCECRKRKGRFIOFKSEKERTKNETETESGNSGYNSDEKTDIASLP
 PKCESSDPEORNSCGPHQADOREPPRAPHSPQASFNLASPERVNTIV"
 BASE COUNT 299 a 302 c 288 g 279 t
 ORIGIN

Query Match 90.2%; Score 1119; DB 6; Length 1168;
 Best Local Similarity 99.1%; Pred. No. 4.7e-255;
 Matches 1125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 106 AGAAGCTGTAGTTGCTGGTGTGTAATGAAGTCTATGAAGGCCCCAGAAATCAACAGT 165
 DB 34 AGAAGCTGTAGTTGCTGGTGTGTAATGAAGTCTATGAAGGCCCCAGAAATCAACAGT 93
 QY 166 CCGTAAGGCTCCAGAGCTCGCTTCACTGCAACCGCTCCAGAGGCTGGAAGCTCATCAT 225
 DB 94 CCGTAAGGCTCCAGAGCTCGCTTCACTGCAACCGCTCCAGAGGCTGGAAGCTCATCAT 153
 QY 226 GTGGGCTCTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 285
 DB 154 GTGGGCTCTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213
 QY 286 TGACCGCTTCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 345
 DB 214 TGACCGCTTCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 273
 QY 346 CCACATGTGAGGCCAGTGAATGGGGAACATGATGATGATGATGATGATGATGATGATGAT 405
 DB 274 CCACATGTGAGGCCAGTGAATGGGGAACATGATGATGATGATGATGATGATGATGATGAT 333
 QY 406 GCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 465
 DB 334 GCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 393
 QY 466 TCTTGTAGTCTGAGTGAATGAACCTTGAAGTGAACCTTGAAGTGAACCTTGAAGTGAACCT 525
 DB 394 TCTTGTAGTCTGAGTGAATGAACCTTGAAGTGAACCTTGAAGTGAACCTTGAAGTGAACCT 453
 QY 526 GCTCCCGGATATTTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
 DB 454 GCTCCCGGATATTTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
 QY 586 TCCGAGGCCAGCGACCTTCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 645
 DB 514 TCCGAGGCCAGCGACCTTCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 573
 QY 646 TGGGACTTTGACTTGGCGGTGCTACCTGGAAGAGCTGGAAGGCCCGCAAGTGTGCAACTGT 705
 DB 574 TGGGACTTTGACTTGGCGGTGCTACCTGGAAGAGCTGGAAGGCCCGCAAGTGTGCAACTGT 633
 QY 706 AAATCTCACTGTGATGGGTGCTCCCAAGACACTGGAAGTGTATTAATATTCAGAGTGT 765
 DB 634 AAATCTCACTGTGATGGGTGCTCCCAAGACACTGGAAGTGTATTAATATTCAGAGTGT 693
 QY 766 ATTATCAAGTTTACCGAGTTTGGTTTTCATTGCTACTTGGGGCAAGTTGGACTTGG 825

FEATURES
source
Location/Qualifiers
1..2051
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE6017788"
/tissue_type="placenta"
/clone_id="PLACE6"
/note="cloning vector: PME18SFL3"

BASE COUNT 580 a 463 c 474 g 534 t

ORIGIN

Query Match 90.8%; Score 1126.4; DB 9; Length 2051;
Best Local Similarity 99.5%; Pred. No. 8.4e-257;
Matches 1130; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

105 CAGAGCGTGAAGTCTGGTCTGGTATGATGAGTCATGAAAGGCCCCAGATGCAACAG 164
192 CAGTGTGAGACGTTCTGGTGTGTATGATGATGATGAAAGGCCCCAGATGCAACAG 251
165 TCCTGAAGGGCTCCAGGCTGCTTCAACATGACCCCTCCAGGGCTGAAGCTCATCA 224
252 TCCTGAAGGGCTCCAGGCTGCTTCAACATGACCCCTCCAGGGCTGAAGCTCATCA 311
225 TGTGGGCTCTCAGTACATGTTGTTGCTTAAGCGTCAGGCCCATGAGCCCATCATCA 284
312 TGTGGGCTCTCAGTACATGTTGTTGCTTAAGCGTCAGGCCCATGAGCCCATCATCA 371
285 ATGACCGGCTTCACTCTCAGAGTACGACGAGGGGGGAGACTTCACCTCGAGATGATCA 344
372 ATGACCGGCTTCACTCTCAGAGTACGACGAGGGGGGAGACTTCACCTCGAGATGATCA 431
345 TCACAAATGTGAGCCAGTATTCGGGGAAATCATGATGACGCTCCAGAACAGTCCG 404
432 TCACAAATGTGAGCCAGTATTCGGGGAAATCATGATGACGCTCCAGAACAGTCCG 491
405 TGCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 464
492 TGCATGATCTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 551
465 ATCTGTGATGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
552 ATCTGTGATGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
525 GGCCTCCCGGATATTTCTGAGAGTCTGCTCTCTGCTGAGCCATTCAGATATTTT 584
612 GGCCTCCCGGATATTTCTGAGAGTCTGCTCTCTGCTGAGCCATTCAGATATTTT 671
585 TTCCGAGAGCCAGGACCTTCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 644
672 TTCCGAGAGCCAGGACCTTCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 731
645 ATGGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 704
732 ATGGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 791
705 TAAATCTCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
792 TAAATCTCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
765 TATATCAAGTTTACCGAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 824
852 TATATCAAGTTTACCGAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 911
825 GACTGAGAGGACCATGCTTCTGAGAGCGGAGCTGATCTTCAATATGAGTCTGCTGCT 884
912 GACTGAGAGGACCATGCTTCTGAGAGCGGAGCTGATCTTCAATATGAGTCTGCTGCT 971
885 GCGGCGGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
972 GCGGCGGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
945 AAGAGGATTTCTGATTCATTTCAAAAGAAATCTGAAAAAGAGAAAGAAAGAAAGAA 1004

Db 1032 AAGAGGATTTCTGATTCATTTCAAAAGAAATCTGAAAAAGAGAAAGAAAGAA 1091
Qy 1005 CTGAGACAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1064
Db 1092 CTGAGACAGAAATGGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1151
Qy 1065 AAACCGGCTTCTCCCTCCCAAAATCTGTAATGATGATGATGATGATGATGATGATGAT 1124
Db 1152 AAACCGGCTTCTCCCTCCCAAAATCTGTAATGATGATGATGATGATGATGATGATGAT 1211
Qy 1125 GCTGTGGCCCTCTCAGACAGGCGGCTGATGATGATGATGATGATGATGATGATGAT 1184
Db 1212 GCTGTGGCCCTCTCAGACAGGCGGCTGATGATGATGATGATGATGATGATGATGAT 1271
Qy 1185 AGGCTCTTTTATTCAGGAGTCTGAGAGGAGTCAATACACTGATGATGATGATGAT 1240
Db 1272 AGGCTCTTTTATTCAGGAGTCTGAGAGGAGTCAATACACTGATGATGATGATGAT 1327

RESULT 3
AX380396 1175 bp DNA linear PAT 18-MAR-2002
LOCUS
DEFINITION
ACCESSION AX380396
VERSION AX380396.1 GI:19575326
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1
AUTHORS Welch A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 1 03-JAN-2002;
Amgen, Inc. (US)

FEATURES
source
Location/Qualifiers
1..1175
/organism="Homo sapiens"
/db_xref="taxon:9606"
27..1175
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28216.1"
/db_xref="GI:19575327"
/translation="MGIVFLHSGSGNEVIEPONATVILKSGARFNCTVSOQWKL
MMALSDMVLYSPMEPIITNDRTSORDGNGFTSEMLIHNVPSDGNIRCSLON
SRHLSGAYITVOYMEELTPSYNLYVAEPEVCLSPSHWTRLPDISMEIGLIVSHS
STFEVPESDLDQSAVSTLALTPQSKNLTICVATWKSLSKRSATVNLVIRPQDTGG
GINIPGVLSLPSLGLFTWKGKVLGAGTMLLPCTLTIRCCCRRCGNCNC
RCQFCRRRGRFRIOFKSEKERNKERTESGNGNSGNSDEQKTDVSLPDKSC
ESSDPEORNSCGPPHQRADQRPSPASHPOASFMILASPEKVSNTTVV"

BASE COUNT 295 a 307 c 286 g 287 t

ORIGIN

Query Match 90.6%; Score 1123.4; DB 6; Length 1175;
Best Local Similarity 99.9%; Pred. No. 4.3e-256;
Matches 1124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 116 GCTTCTGGTCTGATATGAGTCAATGAGAGGCCCCCAAGATGCAACACTCTCGAAGGCC 175
Db 51 GCTTCTGGTCTGATATGAGTCAATGAGAGGCCCCCAAGATGCAACACTCTCGAAGGCC 110
Qy 176 TCCAGAGCTGCTTCAACTGACCGTCTCCAGAGGCTGGAAGCTCATCATGAGGCTCTC 235
Db 111 TCCAGAGCTGCTTCAACTGACCGTCTCCAGAGGCTGGAAGCTCATCATGAGGCTCTC 170
Qy 236 AGTGACATGTTGTTGCTTAAGCGTCAAGGCCATGAGAGCCCATCATCAATGACCGCTTC 295
Db 171 AGTGACATGTTGTTGCTTAAGCGTCAAGGCCATGAGAGCCCATCATCAATGACCGCTTC 230
Qy 296 ACCTCTCAGAGTGAACAGGAGGCGGAACTTCACTCGAGATGATCATCCAAATGTG 355
Db 231 ACCTCTCAGAGTGAACAGGAGGCGGAACTTCACTCGAGATGATCATCCAAATGTG 290

Angen, Inc. (US)
Location/Qualifiers
1. 1240
/organism="Homo sapiens"
/db_xref="taxon:9606"
80.1240
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28218.1"
/db_xref="GI:19575331"
/translation="MERHLLTPPEAVNGSGNVEISPNATVYKSGQARFNTVSOG
WKLIMALSDMYLVSRPMEPIITNDPFIQSDQGNFTSEMIINVEPSDGNIR
SLONRSLHSAVILVQVMEFLIPSVNLVAENPEVETCLPSHMTLDPIDISWELIL
VSHSYTFEPEPSDLSAVSIALITPOSNGTICAVTMSLKARKSAVALVIRCPQ
DTGGGINIPGVLSLPSLGFSLPTWGVGLAGLMILPTCCLTIRCCCCRRCCGC
NCCRCFCRRKRGFRIOFKSEKTEKTEETSGNNSGNSGNDDEKTEETASLP
PKCESSDPQRRNSGCPHQRADORPRASHPOASFNLASPEKVSNTTV"

BASE COUNT 318 a 319 c 305 g 298 t
ORIGIN

Query Match 100.0%; Score 1240; DB 6; Length 1240;
Best Local Similarity 100.0%; Pred. No. 8.7e-284;
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTGTGATCCAGCAACAGTGTGATCACTTCTTACGCTGCAATACAAACACCAT 60
DB 1 AGGTGTGATCCAGCAACAGTGTGATCACTTCTTACGCTGCAATACAAACACCAT 60
QY 61 AACCTGTGCTTGAACATGAAAGGCAATTTGCTCAGCGTTCCAGAGCTGTAGTTC 120
DB 61 AACCTGTGCTTGAACATGAAAGGCAATTTGCTCAGCGTTCCAGAGCTGTAGTTC 120
QY 121 TGGGTCTGTATATAGTATAGAGCCGCCAGATGCAACAGTCTGAGAGGCTTCCA 180
DB 121 TGGGTCTGTATATAGTATAGAGCCGCCAGATGCAACAGTCTGAGAGGCTTCCA 180
QY 181 GGCCTGCTTCACTGCAACCGCTCCCAAGGGCTGGAAGCTATCATGTGGGCTCTCA 240
DB 181 GGCCTGCTTCACTGCAACCGCTCCCAAGGGCTGGAAGCTATCATGTGGGCTCTCA 240
QY 241 CATGTGTGCTTAAAGCTCAGGCGCCATGAGCCCATCATCAATGACCGCTTCACT 300
DB 241 CATGTGTGCTTAAAGCTCAGGCGCCATGAGCCCATCATCAATGACCGCTTCACT 300
QY 301 TCAGAGGTAGACAGGGCGGGAACCTTCACTCGGAGATATATCCACATGTGAGCC 360
DB 301 TCAGAGGTAGACAGGGCGGGAACCTTCACTCGGAGATATATCCACATGTGAGCC 360
QY 361 CAGTATTCGGGGAACATCAGATGACGCTCCAGAACAGTGGCTGATGATCTGCTTA 420
DB 361 CAGTATTCGGGGAACATCAGATGACGCTCCAGAACAGTGGCTGATGATCTGCTTA 420
QY 421 CCTTACCGTCCAAAGTATGGAAGAGCTTTCATTCAGTGTATCTGTAGTCTGCTGA 480
DB 421 CCTTACCGTCCAAAGTATGGAAGAGCTTTCATTCAGTGTATCTGTAGTCTGCTGA 480
QY 481 GAATGAACCTTGAATTAATCTGTCTACCTGACCTGACCTGACCTGACCTGACCTG 540
DB 481 GAATGAACCTTGAATTAATCTGTCTACCTGACCTGACCTGACCTGACCTGACCTG 540
QY 541 CTGGAGACTGCTGCTCTGCTCAGCATTCATTCATTTATTTGTTCCGAGCCAGCA 600
DB 541 CTGGAGACTGCTGCTCTGCTCAGCATTCATTCATTTATTTGTTCCGAGCCAGCA 600
QY 601 CCTTCAAGAGTGAAGAGCTGCTGCTGACCCCAAGAGCAATGAGCACTTGTGACTTG 660
DB 601 CCTTCAAGAGTGAAGAGCTGCTGCTGACCCCAAGAGCAATGAGCACTTGTGACTTG 660
QY 661 CGTGTACTCTGAGAGAGCTGGAAGGCGGCAAGTCTGCAACGTAAATCTCACTGTAT 720
DB 661 CGTGTACTCTGAGAGAGCTGGAAGGCGGCAAGTCTGCAACGTAAATCTCACTGTAT 720
QY 721 TCGGTGTCCCAAGACACTGAGAGTGTATTAATATTCAGGTGTATTAATCAAGTTTACC 780

DB 721 TCGGTGTCCCAAGACACTGAGAGTGTATTAATATTCAGGTGTATTAATCAAGTTTACC 780
QY 781 GACTTTAGTGTATTTTCAATGCTTCTGAGGCAAGTGTGAGTGTGAGTGTGAGTGTGAG 840
DB 781 GACTTTAGTGTATTTTCAATGCTTCTGAGGCAAGTGTGAGTGTGAGTGTGAGTGTGAG 840
QY 841 GCTTTCAGAGCGGAGCTGATACCTTCAATAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 841 GCTTTCAGAGCGGAGCTGATCTTCAATAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 TGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 TGGCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 TCAATTTCAAAAGAAATCTGAAAAGAGAGAACAAAGAAAGAAAGAAAGAAAGAAAGTGG 1020
DB 961 TCAATTTCAAAAGAAATCTGAAAAGAGAGAACAAAGAAAGAAAGAAAGAAAGTGG 1020
QY 1021 AATGAAAACCTCGGCTTCAATTCAGATGACAAAAGACAGAAACCGCTTCTCTCC 1080
DB 1021 AATGAAAACCTCGGCTTCAATTCAGATGACAAAAGACAGAAACCGCTTCTCTCC 1080
QY 1081 TCCCAATCTCTGTAATCCAGTATCTGTAACAAAGAAACAGTATGCTGCTCTCTCA 1140
DB 1081 TCCCAATCTCTGTAATCCAGTATCTGTAACAAAGAAACAGTATGCTGCTCTCTCA 1140
QY 1141 CCAGCGGCTGATCAACGCTCCACAGGCGGAGCAAGTATCCACAGGCTTCTTTAATCT 1200
DB 1141 CCAGCGGCTGATCAACGCTCCACAGGCGGAGCAAGTATCCACAGGCTTCTTTAATCT 1200
QY 1201 GGCAGTCTCTGAGAGTGTAGTATATCAACTGTATATAG 1240
DB 1201 GGCAGTCTCTGAGAGTGTAGTATATCAACTGTATATAG 1240

RESULT 2
AK092516
LOCUS
DEFINITION
Homo sapiens cDNA FLJ35197 fls, clone PLACE6017788, highly similar to IGFS5.
ACCESSION
AK092516
VERSION
AK092516.1 GI:21751130
KEYWORDS
Oligo capping; fls (full insert sequence).
SOURCE
Homo sapiens Placenta cDNA to mRNA, clone_11b.PLACE6
clone:PLACE6017788.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Waga, S., M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Negai, K., and Isogai, T.
TITLE
NEDO human cDNA sequencing project
JOURNAL
unpublished
REFERENCE
2 (bases 1 to 2051)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 23:25:03 ; Search time 3424.45 Seconds

(without alignments)
10538.186 Million cell updates/sec

Title: US-09-729-264-5

Perfect score: 1240
Sequence: 1 aggtgtgagtcacgacacaa.....gtaatacactgtatag 1240

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

GenBank: *
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1240	AX380400	AX380400 Sequence
2	1126.4	90.8	2051	AK092516	AK092516 Homo sapi
3	1123.4	90.6	1175	AX380396	AX380396 Sequence
4	1119	90.2	1168	AX380398	AX380398 Sequence
5	1064	85.8	1139	AX380402	AX380402 Sequence
6	380.6	30.7	1195	AX380404	AX380404 Sequence
7	318.6	25.7	142742	AF121782	AF121782 Homo sapi
8	318.6	25.7	340000	HS210080	HS210080 Homo sapi
9	187.4	15.1	196900	AC020851	AC020851 Mus muscu
10	185.6	15.0	754	AX380408	AX380408 Sequence
11	185.6	15.0	1556	AX380406	AX380406 Sequence
12	185.6	15.0	1556	BC004806	BC004806 Mus muscu
13	176.4	14.2	175861	AC120346	AC120346 Mus muscu
14	119	9.6	140205	AF045450	AF045450 Homo sapi
15	119	9.6	170121	AF064860	AF064860 Homo sapi
16	100	8.1	199665	AF064857	AF064857 Homo sapi
17	58.4	4.7	143266	AC067790	AC067790 Homo sapi
18	58.4	4.7	152276	AC112649	AC112649 Homo sapi
19	58.4	4.7	186968	AC007917	AC007917 Homo sapi
20	58.2	4.7	175861	AC120346	AC120346 Mus muscu
21	58.2	4.7	196900	AC020851	AC020851 Mus muscu
22	56.6	4.6	1427	HS063332	HS063332 Human super
23	56.2	4.5	163295	AC078808	AC078808 Homo sapi
24	56.2	4.5	205933	AC117381	AC117381 Homo sapi
25	55.4	4.5	53908	AC120145	AC120145 Mus muscu
26	55	4.4	1061	HSCAGCTG	AJ000501 Homo sapi
27	55	4.4	123576	AC015969	AC015969 Homo sapi
28	55	4.4	206059	AC127383	AC127383 Homo sapi
29	54.2	4.4	138000	AC079127	AC079127 Homo sapi
30	54.2	4.4	157743	AC067887	AC067887 Homo sapi
31	54.2	4.4	183000	AC025134	AC025134 Homo sapi
32	53.8	4.3	63692	AL603750	AL603750 Human DNA
33	53.8	4.3	160541	AL359271	AL359271 Homo sapi
34	53.8	4.3	206187	AC021172	AC021172 Homo sapi
35	53.4	4.3	73925	AP005626	AP005626 Homo sapi
36	53.4	4.3	164239	AC016074	AC016074 Homo sapi
37	53	4.3	178052	AC118358	AC118358 Rattus no
38	52.8	4.3	138685	AC017582	AC017582 Drosophila
39	52.8	4.3	166249	AC093502	AC093502 Drosophila
40	52.8	4.3	185404	AC104703	AC104703 Drosophila
41	52.8	4.3	341319	AE003536	AE003536 Drosophila
42	52.6	4.2	143501	AC104383	AC104383 Homo sapi
43	52.6	4.2	146836	AC019317	AC019317 Homo sapi
44	52.4	4.2	44768	AL589649	AL589649 Human DNA
45	52.4	4.2	166357	AC091559	AC091559 Homo sapi

ALIGNMENTS

RESULT 1
AX380400
LOCUS AX380400 1240 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 5 from Patent WO0200710.
ACCESSION AX380400
VERSION AX380400.1 GI:19575330
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1. Welcher, A.A., Sarmiento, U.M., Schultz, H.J. and Chute, H.T.
AUTHORS
TITLE B7-like molecules and uses thereof
JOURNAL Patent: WO 0200710-A 5 03-JAN-2002.

